

Every Match	99.5%;	Score 3673.5;	DB 22;	Length 705;
Best Local Similarity	97.6%;	Pred. No. 0;		
Matches 688; Conservative	0;	Mismatches	0;	Indels 17; Gaps 1

RESULT	3
AA004956	
ID	AA004956 standard; Protein: 705 AA.
XX	
AC	AA004956;
XX	
DT	24-OCT-2001 (first entry)
XX	
DE	Human Interleukin 17 receptor; IL-17RH2
XX	
KM	Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist; PRO200040; DNA 164625-2890; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; diabetes mellitus; allergic disease; asthma; demyelinating disease; degenerative cartilaginous disorder; transplantation associated disease
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	1..20
FT	/label= signal_peptide
FT	21..705
FT	/label= Mature_IL_17RH2
FT	107..112
FT	/note= "N-myristoylation site"
FT	118..121
FT	/note= "Asn is N-glycosylated"
FT	152..157
FT	/note= "N-myristoylation site"
FT	186..189
FT	/note= "Asn is N-glycosylated"
FT	198..201
FT	/note= "Asn is N-glycosylated"
FT	211..214
FT	/note= "Asn is N-glycosylated"
FT	238..241
FT	/note= "Asn is N-glycosylated"
FT	248..251
FT	/note= "Asn is N-glycosylated"
FT	319..324
FT	/note= "N-myristoylation site"
FT	334..337
FT	/note= "Asn is N-glycosylated"
FT	357..360
FT	/note= "Asn is N-glycosylated"
FT	391..394
FT	/note= "Asn is N-glycosylated"
FT	438..443
FT	/note= "N-myristoylation site"
FT	453..473
FT	/note= "Transmembrane domain"
FT	516..521
FT	/note= "N-myristoylation site"
FT	552..555
FT	/note= "CAMP/GMP-dependent protein kinase phosphorylation site"
FT	583..586
FT	/note= "Glycosaminoglycan attachment site"
FT	612..617
FT	/note= "N-myristoylation site"
FT	692..697
FT	/note= "N-myristoylation site"
FT	696..701
FT	/note= "N-myristoylation site"
FT	700..705
FT	/note= "N-myristoylation site"
XX	
PN	NO200146420-A2.
XX	
DD	28-JUN-2001.
XX	

Db	241	spqitlthntdtpvclciqvwpdepdsvrtnlopfredprahgnlwgaarlrlltqlgswl	3000
Qy	301	LDAPCSLPAEALCALCRAAGSGPPCPPLVPLPSMENTVD-----VNSSE	3433
Db	301	ldapcslpaeealcvrapsgpccplvplswentvdkvlefrlkhgnplcvgnasse	3600
Qy	344	KLQLOECIMADPSLGLKDDVLLLEFRGQDNRSLCALEPSCGCTSLPSKASRRARIGEYL	4030
Db	361	klqgeclwadsjgplkadvlllefrgqdnrsllcalepsgctslpskasttraarlgelyl	4200
Qy	404	LDLQSGOCLQMDDDLQALNACPMQKYLHRRNALWLNACLLFPAALSLILLEKKDRAGK	4630
Db	421	ldqlsggcqlqwdddlgalwacpmqkyihkrvalwlacllffaalsllllkkhknag	4800
Qy	464	WRLRLKQVRSAGAAARGAALLLTYSADDSGFERLVGALASALCOLPLRAVADLMSRREL	5230
Db	481	wrlrkqvrsagaargraallllysaddsgferlvgalasalcolplrvavdlwstrrels	5400
Qy	524	AQGPAPWHAQRROTLDGCGVTVLLFSGANALCSEWLODGVSPGAHGPHDAFRASLSC	5830
Db	541	aqgpapwhaqrrotldgcvtvllfsgavalcsewldgvsqgahgphdafraasisc	6000
Qy	584	VLPDLQGRAPGSVYGAFCEDFLTPDAVPALFRVVPFTLPSOLPDLGLALQOPRAPRSG	6430
Db	601	vlpdlqgrapgsyvgafcdfltpdavpalfrvvpftlpsqldpdlgalqgraprsq	6600
Qy	644	RLOERAEQVSRALQPALDSYHPHGPFPAPRGVPGAGPGAGDGT	688
Db	661	rlqeraeqvstralqpaldsyhpqgtpaprgvpgagpgagdgdt	705
RESULT 4			
AA87606	ID AAB87606 standard; Protein: 705 AA.		
XX	AA87606;		
XX	15-MAY-2001 (first entry)		
XX	Human PRO20040.		
XX	Human; PRO protein; mapping.		
XX	Homo sapiens.		
XX	W0200116318-A2.		
XX	08-MAR-2001.		
XX	24-AUG-2000; 2000MO-US23328.		
XX	01-SEP-1999; 99MO-US20111.		
XX	15-SEP-1999; 99MO-US21090.		
XX	07-DEC-1999; 99US-0169495.		
XX	09-DEC-1999; 99US-0170262.		
XX	11-JAN-2000; 2000US-0175481.		
XX	18-FEB-2000; 2000MO-US04341.		
XX	18-FEB-2000; 2000MO-US04342.		
XX	22-FEB-2000; 2000MO-US04414.		
XX	01-MAR-2000; 2000MO-US05601.		
XX	03-MAR-2000; 2000US-0187202.		
XX	25-APR-2000; 2000US-0199397.		
XX	PR 22-MAY-2000; 2000MO-US14042.		
XX	05-JUN-2000; 2000US-0209832.		
XX	(GETH ) GENENTECH INC.		
XX	Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ,		
XX	PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WT.		
XX	WPI; 2001-183260/18.		
XX	N-PSDB; AAF92138.		

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 17:01:01 ; Search time 43.62 Seconds

(without alignments)  
610.707 Million cell updates/sec

Title: US-09-608-918-10

Sequence: 1 MPVWFLSLALGRSPVLS.....TPAPGRGVGPGAGGAGDGT 688

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128.5	3.5	1711	1	PTPO_RAT
2	126	3.4	3707	1	PGBM_MOUSE
3	124.5	3.4	761	1	POOF_KLEPN
4	113.5	3.1	587	1	NDC2_RAT
5	109.5	3.0	428	1	NER3_HUMAN
6	108.5	2.9	753	1	U06_HSVB
7	107	2.9	825	1	IL4R_HUMAN
8	106.5	2.9	1709	1	SM_HUMAN
9	105.5	2.9	741	1	TLE4_RAT
10	105	2.8	766	1	TLE4_HUMAN
11	102	2.8	1208	1	RCO4_HUMAN
12	101.5	2.7	418	1	NER3_RAT
13	100	2.7	1011	1	MAK6_HUMAN
14	99	2.7	505	1	CXAA_MOUSE
15	99	2.7	922	1	W70T_MOUSE
16	98.5	2.7	890	1	TYO3_HUMAN
17	98	2.7	5430	1	ACE7_HUMAN
18	97.5	2.6	676	1	U06_HSV1
19	97.5	2.6	850	1	NRG2_HUMAN
20	97.5	2.6	897	1	CYRB_HUMAN
21	97	2.6	993	1	AFSR_STRCO
22	96.5	2.6	992	1	SNX1_HUMAN
23	95	2.6	528	1	VRP3_GIALA
24	95	2.6	577	1	TIE4_MOUSE
25	94.5	2.6	209	1	CXX1_HUMAN
26	94	2.5	392	1	GAG_BLYAV
27	94	2.5	2205	1	POLN_RUBYT
28	94	2.5	4393	1	PGBM_HUMAN
29	93.5	2.5	459	1	CAH9_HUMAN
30	93.5	2.5	775	1	POOF_PSEAE
31	93.5	2.5	880	1	TYO3_RAT
32	93.5	2.5	1442	1	PTC1_CAEEL
33	92.5	2.5	379	1	YP85_CAEEL

34	92.5	2.5	537	1	ICAI_MOUSE	P13597 mus musculu
35	92.5	2.5	867	1	SSPO_BOVIN	P98167 bos taurus
36	92.5	2.5	947	1	MAKE_HUMAN	O99558 homo sapien
37	92.5	2.5	1473	1	NALI_HUMAN	O9C000 homo sapien
38	92	2.5	443	1	W70T_RAT	O35828 rattus norv
39	92	2.5	492	1	CPSL_PIG	P15540 sus scrofa
40	92	2.5	553	1	MIS_RAT	P49000 rattus norv
41	92	2.5	672	1	NOO3_PARDE	P29915 paracoccus
42	92	2.5	843	1	Y096_CAEEL	P41846 caenorhabd1
43	92	2.5	1705	1	PTPO_MOUSE	P70289 mus musculu
44	90.5	2.5	368	1	GALT_HUMAN	O60755 homo sapien
45	90.5	2.5	508	1	EPOR_HUMAN	P19235 homo sapien

## ALIGNMENTS

RESULT	ID	PTPO_RAT	STANDARD	PRT	1711 AA.
AC	064612				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-JUN-1999 (Rel. 38, Last annotation update)				
DE	Osteosticellular protein tyrosine phosphatase precursor (EC 3.1.3.48)				
DE	(OST-PTP).				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Osteosarcoma;				
RX	MEDLINE=95074080; PubMed=7527035;				
RA	Mauro L.J., Olmsted E.A., Skrobacz B.M., Mourey R.J., Davis A.R.,				
RA	Dixon J.E.;				
RT	"Identification of a hormonally regulated protein tyrosine				
RT	phosphatase associated with bone and testicular differentiation.";				
RL	J. Biol. Chem. 269:30659-30667(1994).				
CC	-1- FUNCTION: MAY FUNCTION IN SIGNALING PATHWAYS DURING BONE				
CC	REMODELING, AS WELL AS SERVE A BROADER ROLE IN CELL INTERACTIONS				
CC	ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS. OPTIMAL PH FOR				
CC	PHOSPHATASE ACTIVITY IS 5.6.				
CC	-1- ASSOCIATED WITH DIFFERENTIATION IN BONE AND TESTIS.				
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein				
CC	tyrosine + phosphate.				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS: A PRESUMED ALTERNATE TRANSCRIPT OF 4.8-5.0				
CC	KILOBASES, WHICH MAY LACK PTP DOMAINS, IS PRESENT IN PROLIFERATING				
CC	OSTEOBLASTS, BUT NOT DETECTABLE AT OTHER STAGES.				
CC	-1- TISSUE SPECIFICITY: BONE AND TESTIS. IN THE LATTER, RESTRICTED TO				
CC	THE BASAL PORTION OF THE SEMINIFEROUS TUBULE.				
CC	-1- DEVELOPMENTAL STAGE: UP-REGULATED IN DIFFERENTIATING CULTURES OF				
CC	PRIMARY OSTEOBLASTS AND DOWN-REGULATED IN LATE STAGE MINERALIZING				
CC	CULTURES. IN TESTIS, EXPRESSION IS HIGHEST BETWEEN STAGES I AND				
CC	VI WHEN MATURING SPERMATIDS REMAIN BURIED WITHIN THE SERTOLI				
CC	EPITHELIUM.				
CC	-1- INDUCTION: BY PARATHYROID HORMONE AND CYCLIC AMP ANALOGS.				
CC	-1- PTM: THE CYTOPLASMIC DOMAIN CONTAINS POTENTIAL PHOSPHORYLATION				
CC	SITES.				
CC	-1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.				
CC	-1- SIMILARITY: CONTAINS 10 FIBRONECTIN TYPE III-LIKE DOMAINS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
DR	EMBL: L36884; AAA63911.1; -				
DR	HSSP: P18052; IYFO.				







```
FT DISULFID 1211 1234 BY SIMILARITY.
FT DISULFID 1237 1246 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1275 1287 BY SIMILARITY.
FT DISULFID 1277 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1363 1372 BY SIMILARITY.
FT DISULFID 1365 1379 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2213 BY SIMILARITY.
FT DISULFID 2268 2313 BY SIMILARITY.
FT DISULFID 2365 2413 BY SIMILARITY.
FT DISULFID 2456 2506 BY SIMILARITY.
FT DISULFID 2554 2599 BY SIMILARITY.
FT DISULFID 2641 2686 BY SIMILARITY.
FT DISULFID 2831 2876 BY SIMILARITY.

Query Match 3.4%; Score 126; DB 1; Length 3707;
Best Local Similarity 19.6%; Pred. No. 0.29;
Matches 97; Conservative 41; Mismatches 111; Indels 246; Gaps 25;

Qy 4 PWFLSLALGRSPVLSRLVGPQDAT-----HCSP-----GLSC----- 39
Db 272 PQFLPSVPGPS-----ACGPQASCHSGCIPRDYLCDDQEDCRDGSDELGCASP 323
Qy 40 -----RLWDSILCLPGDI-----VPAGCPVLAPTHLOTEL 71
Db 324 PCPEPEACENGHCALKWRCD-----GDFCEDRTDEANCSVKQPGVCGPTHFCVST 378
Qy 72 LR-----COKETCDCLRLVAVHLVGHWEPEDEKFGGAADSGVEEPNLSIOAQ 124
Db 379 NRCIPASPHCEDESC-----PDRSEFGCMPPQVVTTPQOQ-SIOA- 418
Qy 125 VVLSFOAYPTARCVLLEVQVPAALVFGOSGVSVYDCEFAALGSEVRWISVTPRYEKE 184
Db 419 -----SRGQ-----TVFTCTVATGVPTPIINW-----R 441
Qy 195 LNHTQQLPALPWLNVASGDNVHLNVSEQHFG-----LSLYW 224
Db 442 LN-WGHIPAPRVMTSEGGRGTLIRDVKEADQAYTCEAMNSRGWFGIPDGVLV- 499
Qy 225 NOVQGPXPRWHKLTGQIITLHNTDLYPCLIQWPLEPDSVRTNICP-----FREDPR 280
Db 500 --PORGCPDGHFYLE-----DSASCLPCFCFG-----TNVCQSSLFRDQIR 541
Qy 281 A---HONLQWQARLRLLTLQSWLLDAPCSLPAEALCWAPGDCPCQPLVPPLSWENVTV 337
Db 542 LSFDPNDFKGVNVTM-----PSQPGVPPPLSSTQLQI 573
Qy 338 D-----VNSSEKLOLQECILWA-----DSLQ-----PL-KD 361
Db 574 DPALQEQFLVLSRRLVHDAFWALPKQPLGNKVDYSYGGFLRVKRYELARGMLEPVQXP 633
Qy 362 DVLL-----LETRG 370
Db 634 DVILVGAGYRLHSRG 648

RESULT 3
POQF_KLEPN
ID POQF_KLEPN STANDARD; PRT; 761 AA.
AC P27508;
DT 01-AUG-1992 (Rel. 23, Created)
```

```
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Coenzyme PQQ synthesis protein F (EC 3.4.99.-).
GN POQF.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 418;
RX MEDLINE=92212293; PubMed=1313537;
RA Meulenberg J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae poq
  operon.";
RL Mol. Gen. Genet. 232:284-294(1992).
CC -I- FUNCTION: REQUIRED FOR COENZYME PYROLO-QUINOLINE-QUINONE (POQ)
CC BIOSYNTHESIS. IT IS THOUGHT THAT THIS PROTEIN IS A PROTEASE THAT
CC CLEAVES PEPTIDES BOND IN A SMALL PEPTIDE (GENE POQA) WHICH WOULD
CC PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH ARE NECESSARY
CC FOR THE SYNTHESIS OF POQ.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58778; CAA41584.1; -
DR PIR; S20458; S20458.
DR MEROPS; M16.006; -
DR InterPro; IPR001431; Peptidase_M16.
DR Pfam; PF00675; Peptidase_M16; 1.
DR PROSITE; PS00143; INSULINASE; 1.
KW POQ; Hydrolase; Metalloprotease; Zinc.
FT METAL 49 49 ZINC (BY SIMILARITY).
FT ACT_SITE 52 52 BY SIMILARITY.
FT METAL 53 53 ZINC (BY SIMILARITY).
FT METAL 130 130 ZINC (BY SIMILARITY).
SQ SEQUENCE 761 AA; 83616 MW; 187182DBC8E839B5 CRC64;

Query Match 3.4%; Score 124.5; DB 1; Length 761;
Best Local Similarity 24.5%; Pred. No. 0.049;
Matches 120; Conservative 47; Mismatches 171; Indels 151; Gaps 27;

Qy 282 HONLQWQARLRLLTLQSWLLDAPCSLPAEALCWAPGG---DPCQPLVPPLSWENVTV 338
Db 182 HRHYVARMQL-----W-LQGQSLALGELAARFAAGLAAGEAPPAPPLRGFTA- 234
Qy 339 VNSSEKLOL-----QECLEWADSLGPKDVLLELTRGPQDNRSICALEPSCGCTSLPSKAST 394
Db 235 -----LQAVSSOPALWRCPLIALSDNVLLR----- 261
Qy 395 RAARLGYLLQD-----LQSGQCILQWDDDLGALWACPMKDYIHKRWLVWLACLLFAAL 450
Db 262 -----EFLDEAPGSLMAGLRQRRLAGDVALNW-----LYODRHLGWLALVVFASDR 307
Qy 451 SLILLKKKHAKGWLRLKQDV-----RSGAAA-----RGRALLYSADDS 492
Db 308 PEEVDROITH---WLQALQOTTPEQQHYTQLSRRLRQALSPLDLQRLQRAFGFAPGAPPA 364
Qy 493 GFERLVGAL-----ASALCO-----LPLRVAVDLWSRRLSAGQGVANFHA 533
Db 365 GFADFCALQVAPSVSLACQTVSPGEPVATGQFSLPL-----AVALCSEWLQDGV- 420
Qy 534 QRRQTL-----QEGGVVVLFFSPG-----AVALCSEWLQDGV- 567
Db 421 QAAGDLVAKCEKAPALLHLPSGCDPRLRLRPFFYCSPOAEGCLARGEQLRPLLAALRH 480
```

```

QY 568 PGARPHDAFRAS--LSCVLPDFLOGRAGSYVACFDRLHPDA--VPALEFTVPEVTL 623
DB 481 AGGHEMHLFDGSMQTLTLPDPE--PCRREPAILOALIRLALFVASTLPS-ESIAIRHL 537
QY 624 PSQULDFGALQOQPR---APRSGRLQERAEQVSRALQPALDSFHHPCGPAPC---RGVG 677
DB 538 MAQLEPRLGTSQHQKGMALAGSNAEDKQWAKOLS-LITAFVNP-WEAPAPCRRGVE 595
QY 678 PGAGPGAGD 686
DB 596 RLVPYR-GD 603

RESULT 4
NDCC2_RAT STANDARD: PRT: 587 AA.
AC P70545:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intestinal sodium/dicarboxylate cotransporter (Na(+)/dicarboxylate cotransporter).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MITAR;
RA MEDLINE=97107437; PubMed=8950177;
RT "Cloning of the cDNA for a rat intestinal Na+/dicarboxylate cotransporter reveals partial sequence homology with a rat intestinal mucin."
RT cotransporter reveals partial sequence homology with a rat intestinal mucin."
RL Biochim. Biophys. Acta 1309:58-62(1996).
CC -1- FUNCTION: COTRANSPORT OF SODIUM AND DICARBOXYLATES SUCH AS SUCCINATE AND CITRATE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE NADC/P/PHO87 FAMILY OF TRANSPORTERS. NADC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U51153; AAB97095.1;
DR InterPro; IPR001898; Na_sulph_symp.
DR Pfam; PF00939; Na_sulph_symp; 3.
DR PROSITE; PS01271; NA_SULFATE; 1.
RW Transport; Transmembrane; Sodium transport; Symport.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
SO SEQUENCE 587 AA; 63531 MW; F9B74F921BDC8712 CRC64;

Query Match 3.1%; Score 113.5; DB 1; Length 587;
Best Local Similarity 21.3%; Pred No. 0.26;
Matches 91; Conservative 47; Mismatches 166; Indels 123; Gaps 20;
QY 312 ALCWRAPGDP-----CQPL-VPLSWENVTVVNSSEKIQLECIAMDSLGLPKRDVL 364

```

```

DB 2 ATCPALMANRFTYILVCLPFLPLPLIVOTKAYCAYSITLALMCTEALPLAIVL 61
QY 365 -----LEETRGPODNRSICALPEPSCGTSLSKASTRAARLGEYLIQDLOSQCQLQWDDDL 420
DB 62 FPIYLFILMGIMD-----ASEGLHRLSRTPI-----YCLSGADGHCCTLEP-- 105
QY 421 GALTACPMADKYIHK-RNALVWLACLLPAAALSL-----TLLKKDHA-KGWLRLIKQ 470
DB 106 -AQTHCPSPASYSRASAALLILGFMVLVYAFLSMISMNTATTAMVPICHAIVLEQLQASK 164
QY 471 DVRSQ-----
DB 165 DVEGGNNPFEELQEECPQKVTKLNGQPVASBPSEPRTQTOEHHRFSGLSCLCYSA 224
QY 490 DDSGFERLYGALASALCOLPLRVAVDLSRRELASQGVAMFHAORROTQEGGVLLF 549
DB 225 SIGIDNLTGTTPLVLVQGVN-----SLFPKQCGELCEFYVFRPLPHDHLAAQ---LAM 278
QY 550 SPGAVALC---SEWL-----QDVGSGPAGHPDAFPAASLCVLPDFLQ 590
DB 279 ATGPLPCQQLPEELMLNGRGRTRACGFPQHDPVQAQA---HEFCREDLFTVLVLY 335
QY 591 ----GRAGSYVG---ACF-----DRLHPDAVPLFRVYPTLPQSLPDTGLAQQPRA 639
DB 336 VLWFTREPGFFPGGDTVFANEGKQSMPSDGTVAIFLSLVFIIIPSKIP--GLMEDPK 392
QY 640 PRSGRLQ 646
DB 393 P--GKTK 397

RESULT 5
NER3_HUMAN STANDARD: PRT: 428 AA.
AC Q9UQA9; Q9NOEL;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Stialidase 3 (EC 3.2.1.18) (Membrane stialidase) (ganglioside stialidase) (N-acetyl-alpha-neuraminidase 3).
GN NEU3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9935353; PubMed=10405317;
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;
RT "Cloning, expression, and chromosomal mapping of a human ganglioside stialidase."
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX PubMed=10861246;
RA Monti E., Bassi M.T., Papini N., Riboni M., Manzoni M., Venerando B., Croci G., Preti A., Ballabio A., Tettamanzi G., Borsani G.;
RT "Identification and expression of NEU3, a novel human stialidase associated to the plasma membrane."
RL Biochem. J. 349:343-351(2000).
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of the lipid bilayer at the level of membrane-bound stialyl glycoconjugates.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-, alpha-(2->8)-glycosidic linkages of terminal sialic residues in oligosaccharides, glycoproteins, glycolipids, colominic acid and synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis, adrenal gland and thymus, followed by pancreas, liver, heart and

```

CC thymus. weakly expressed in kidney, placenta, brain and lung.  
 CC -!- MISCELLANEOUS: Optimum pH is 3.8.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
 CC -!- SIMILARITY: CONTAINS 3 BNR REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB008185; BAA82611.1; -  
 DR EMBL; Y18563; CAB96131.1; ALT\_INIT.  
 DR MIM; 604617; -  
 DR InterPro; IPR002860; BNR.  
 DR Pfam; PF02012; BNR; 3.  
 KW Hydrolase; Glycosidase; Membrane; Repeat.  
 FT REPEAT 129 140 BNR 1.  
 FT REPEAT 203 214 BNR 2.  
 FT REPEAT 254 265 BNR 3.  
 FT SITE 24 27 FRIP MOTIF.  
 FT ACT\_SITE 25 25 BY SIMILARITY.  
 FT ACT\_SITE 45 45 POTENTIAL.  
 FT ACT\_SITE 50 50 POTENTIAL.  
 FT ACT\_SITE 87 87 POTENTIAL.  
 FT ACT\_SITE 225 225 POTENTIAL.  
 FT ACT\_SITE 245 245 POTENTIAL.  
 FT ACT\_SITE 340 340 BY SIMILARITY.  
 FT ACT\_SITE 370 370 POTENTIAL.  
 FT ACT\_SITE 387 387 POTENTIAL.  
 SQ SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 3.0%; Score 109.5; DB 1; Length 428;  
 Best Local Similarity 21.1%; Pred. No. 0.35;  
 Matches 107; Conservative 44; Mismatches 158; Indels 199; Gaps 26;  
 QY 170 EVRIWSTYTOP--RYEKELNHTQOLPALPWL-----NVSADGNVHLYLVNSE 214  
 DB 3 EVTTCFSNPLFRQEDDRGITHYIPALLYPTHTFLAFKRSRTRDEALHLVLR--R 60  
 QY 215 EQHFLSLYWNQVQPKPRWKNKTGTQIITLHNDLVPCLCIQWVPLPDSVRTNCP 274  
 DB 61 GLRIGOLYQW---GELKPLMEATLPG-----HRTMNP-----PWQKSGCVFFLFIC 106  
 QY 275 FREDPRAHONL---WQAAKRLRLTLQSWLLDAPCSLPAAALCWRAPGDPCQPLVPPLS 331  
 DB 107 VRGHVTERQQIVSGRNAARLCFTYSQ-----DAGC-----S 137  
 QY 332 WENV---TVDVNSSEKLOQLCWAD-SLGP-----LKDDVILLETGPGQDNRLCALEP 382  
 DB 138 WSEVRDLTEEVIGSELKH-----WATFVAGPGHGHIQLSGRVLP-----AYTYIIP 184  
 QY 383 SG--CTSLSPSKASTRAARLGEVLLQDLQSGQCQLQWDDDLGALWACPMKYIHKRW--AL 438  
 DB 185 SWFFCFQLPCKTRPHS-----LMYISDDLGVTH-----HGRLIRPM 221  
 QY 439 VWLAC-----LFFAALS-----LILLKKDHAKGWRL----- 467  
 DB 222 VTVEVEAEVTRAGHPVLVYCSARTPNCRABALSTDHGEGFORLALSRQCEPPHCQCG 281  
 QY 468 -----LKDVRSGGAARGAALLLSADDSGF 494  
 DB 282 SVFSRPLPIPHRCODSSSKDAPTQQSGPSGSLREEAGTPE-----SWLLYSHTPSRK 338  
 QY 495 ERLVGALASALCOLPLRVADLWSRELQAQGVPMVFAHQRRQTLQEGGVVLLFSPGAV 554  
 DB 339 QRV--DLGIYLNQTPLEAA--CWSRWIILHCGPCGI---SDLAALKEEGLFGCLFECGTK 391  
 QY 555 ALCSE-----WLQDGVSGPG 569

DB 392 QECEQIAFRLEFTHREILSHLOQDCTSPG 419  
 RESULT 6  
 UL06\_HSVB STANDARD; PRT; 753 AA.  
 AC P28944; 1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Virion gene 56 protein.  
 GN 56.  
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=31520;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92295566; PubMed=1318606;  
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;  
 RL "The DNA sequence of equine herpesvirus-1";  
 RL Virology 189:304-316(1992).  
 CC -!- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA  
 CC PACKAGING.  
 CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,  
 CC EHV-1 56, EBV BBF1, HCMV UL104, AND VZV 54.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M85664; AAB02491.1; -  
 DR PIR; B36801; WZBEE8.  
 DR InterPro; IPR002660; Herpes\_UL6.  
 DR Pfam; PF01763; Herpes\_UL6; 1.  
 DR PRODOM; PD003210; Herpes\_UL6; 1.  
 SQ SEQUENCE 753 AA; 83992 MW; C5E118F78BBED203 CRC64;  
 Query Match 2.9%; Score 108.5; DB 1; Length 753;  
 Best Local Similarity 22.1%; Pred. No. 0.88;  
 Matches 100; Conservative 46; Mismatches 164; Indels 143; Gaps 22;  
 QY 295 TLQSWLLDAPCSLPAAALCWRA--PGGD-----PCOPLVP-----P 329  
 DB 156 TLQSWLLDFVRSITA---CFASSEPDCGTASFAYIDWIACLGLIPLQLKRAPGATVHP 211  
 QY 330 LSWENVTVDVNSSEKLOQLCWADSLGLKDDVLLLETGPGQDNRLS----- 377  
 DB 212 KLMKRLPTDVPES-----LESCV-----DERDLIAGKLYVANSLLR 245  
 QY 378 ---CALEPSGCT-SLPSKASTRAARLGEVLLQDL-----OSGOCLOLWD---DDLQ 421  
 DB 246 EGLEAVVELACTASVIMDYDRVNIFFHYTRREVAIDSTTGKRGECCLVWQPIWKGDS 305  
 QY 422 ALWACPMKY-----IHKRWALVWLACLLFFAALS-----LILLKKDHAKGWRLKQDVRSG 475  
 DB 306 VLFDSPLQICGECVNCCHALREHAKLCOLLNTPVKILVGRKKDEAQG-----PG 355  
 QY 476 AAARGRAALLYSADDSGFERLVGALASALCOLPLRVADLWSRELQAQGVPMVFAHQ 535  
 DB 356 WASKAVDKLM-----GEHEELHSSSAASRL---VKLVNKMMSMRHIGDITETVRSYLINE 406  
 QY 536 RQTLQEGGVVLLFSPGVALCSEWLQDGVSGPGAHGP-HDAFRASLSCVLPDFLOGRAP 594  
 DB 407 TSTNLLSGAQVDTSLPG-----FGSGKTKQGNMVPQEAFTSVINGINMGLEG--- 456  
 QY 595 GSYVGACFDRLLLHPDAVPALFTVPVFTLPQSLPDFLQALQOPRAPRSGRLQERAEQVSR 654

DB 457 --YVN-----NLKFTIEDLRGNS-----GLLDQLR-DESEITHLREQLLR 495  
OY 655 ALOPALDSYFHPGPFA--PGRCVGPAGAG 685  
DB 496 VSOAADGSTOPGASSALPGSGAKSGAG-GLG 527

RESULT 7  
IL4R\_HUMAN STANDARD: PRT; 825 AA.  
AC P24394;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Interleukin-4 receptor alpha chain precursor (IL-4R-alpha) (CD124 antigen).  
GN IL4R OR IL4RA OR 582J2.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Peripheral blood;  
RX MEDLINE-90171849; PubMed-23079314;  
RA Idzerda R.L., March C.J., Mosley B., Lyman S.D., Bos T.V., Gimpel S.D., Din W.S., Grabstein K.H., Widmer M.B., Park L.S., Cosman D., Beckmann M.P.;  
RT "Human interleukin 4 receptor confers biological responsiveness and defines a novel receptor superfamily.";  
RL J. Exp. Med. 171:861-873(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91120547; PubMed-2278997;  
RA Gallizzi J.-P., Zuber C.E., Harada N., Gorman D.M., Djossou O., Kastlein R., Banchereau J., Howard M., Miyajima A.;  
RT "Molecular cloning of a cDNA encoding the human interleukin 4 receptor.";  
RL Int. Immunol. 2:669-675(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99425270; PubMed-10493829;  
RA Lofthus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhman J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-D., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";  
RL Genomics 60:295-308(1999).  
RN [4]  
RP VARIANTS VAL-75; ALA-400; ARG-431; LEU-436 AND PRO-786.  
RX MEDLINE-97224413; PubMed-9070874;  
RA Delchmann K., Bardutzky J., Forster J., Heinemann A., Kuehr J.;  
RT "Common polymorphisms in the coding part of the IL4-receptor gene.";  
RL Biochem. Biophys. Res. Commun. 231:696-697(1997).  
RN [5]  
RP VARIANT ATOPIC ARG-576.  
RX MEDLINE-98041803; PubMed-9392697;  
RA Hershey G.K.K., Friedrich M.F., Esswein L.A., Thomas M.L., Chaila T.A.;  
RT "The association of atopy with a gain-of-function mutation in the alpha subunit of the interleukin-4 receptor.";  
RL New Engl. J. Med. 337:1720-1725(1997).  
RN [6]  
RP VARIANT ATOPIC ASTHMA VAL-75.  
RX MEDLINE-96282087; PubMed-9620765;  
RA Mitsuysu H., Izuhara K., Mao X.-Q., Gao P.S., Arinobu Y., Enomoto T., Kawai M., Sasaki S., Dake Y., Hamasaki N., Shirakawa T., Hopkin J.M.;  
RT "The50val variant of IL4R alpha upregulates Ige synthesis and associates with atopic asthma.";  
RL Nat. Genet. 19:119-120(1998).  
RN [7]  
RP VARIANT ATOPIC ASTHMA VAL-75.

RX MEDLINE-99322293; PubMed-10390422;  
RA Noguchi E., Shimasaki M., Arinami T., Takeda K., Yokouchi Y., Kobayashi K., Imoto N., Nakahara S., Matsui A., Hamaguchi H.;  
RT "No association between atopy/asthma and the Ile50val polymorphism of IL-4 receptor.";  
RL Am. J. Respir. Crit. Care Med. 160:342-345(1999).  
RN [8]  
RP VARIANTS PRO-503 AND ARG-576.  
RX MEDLINE-99250314; PubMed-10233717;  
RA Kruse S., Japha T., Tedner M., Sparholt S.H., Forster J., Kuehr J., Delchmann K.A.;  
RT "The polymorphisms S503P and Q576R in the interleukin-4 receptor alpha gene are associated with atopy and influence the signal transduction.";  
RL Immunology 96:365-371(1999).  
RN [9]  
RP VARIANT ALA-752.  
RX MEDLINE-20143377; PubMed-10677312;  
RA Ober C., Leavitt S.A., Tsalenko A., Howard T.D., Hokl D.M., Daniel R., Newman D.L., Wu X., Parry R., Lester L.A., Solway J., Blumenthal M., King R.A., Xu J., Meyers D.A., Blecker E.R., Cox N.J.;  
RT "Variation in the interleukin 4-receptor alpha gene confers susceptibility to asthma and atopy in ethnically diverse populations.";  
RL Am. J. Hum. Genet. 66:517-526(2000).  
RN [10]  
RP VARIANT ATOPIC ARG-576.  
RX MEDLINE-20269830; PubMed-10809862;  
RA Olso N., Fukui K., Ishii M.;  
RT "Interleukin 4 receptor alpha chain polymorphism Gln551Arg is associated with adult atopic dermatitis in Japan.";  
RL Br. J. Dermatol. 142:1003-1006(2000).  
RN [11]  
RP VARIANT PRO-786.  
RX MEDLINE-21405389; PubMed-11513543;  
RA Andrews R.P., Burrell L., Rosa-Rosa L., Cunningham C.M., Brzezinski J.L., Bernstein J.A., Khurana Hershey G.K.;  
RT "Analysis of the Ser786Pro interleukin 4 receptor alpha allelic variant in allergic and nonallergic asthma and its functional consequences.";  
RL Clin. Immunol. 100:298-304(2001).  
RN [12]  
RP VARIANT ILE-579.  
RX MEDLINE-21182061; PubMed-11285129;  
RA Lozano F., Placés L., Vila J.M., Padilla O., Arman M., Gimferrer I., Sazano B., Lopez de la Iglesia A., Miserachs N., Vives J.;  
RT "Identification of a novel single-nucleotide polymorphism (Val554Ile) and definition of eight common alleles for human IL4RA exon 11.";  
RL Tissue Antigens 57:216-220(2001).  
CC -1- FUNCTION: THIS IS A RECEPTOR FOR INTERLEUKIN-4. A SOLUBLE FORM OF THE IL-4 RECEPTOR MAY REPRESENT A REGULATORY MOLECULE SPECIFIC FOR IL-4-DEPENDENT IMMUNE RESPONSES.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A COMMON GAMMA CHAIN.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- DISEASE: DEFECTS IN IL4R ARE A CAUSE OF ATOPIC ASTHMA.  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD124 entry;  
CC WWW=<http://www.ncbi.nlm.nih.gov/prov/cd/cd124.htm>  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X52425; CA36672.1; -  
CC EMBL: AC004525; AAC23495.1; -  
CC PIR: A60386; A60386;  
CC PIR: A47603; A47603;  
CC MIM: 147781; -  
CC Interpro: IPR002996; CRIA.





```

Db 1301 -----LOEGPAASLSFLVATRAHAGAYSCQAQDAQGTSSRPAALQVLYAPQDAVLSSFRD 1356
QY 500 ALASALCOLPLRVAVDLWSRRRLSAGCPVAVFPAQRQRTLQEGGVVLLFSPGAVALCSE 559
Db 1357 SRARSNAVLCQTVDSPPPALASHDGKV-----LATSGVHSLASGTHVQVARN 1407
QY 560 WL-----QDGVSG-----PGAHGPHDAFRASISCV 584
Db 1408 ALRLQVQDVPAGDDTVCTAQNLLSGISITIGRLQVEGARVVAEPGLDVPEGA-ALNLSR 1466
QY 585 LPDFLQGRAP-GSYGACF--DRLLHPDAVPALFRT-----VPVFTLPSQLPDFLQALQ 635
Db 1467 L---LGGPGVGNSTFAWVNDRLRAEPVPTLAFTHVARAQAGMYHCLAEPL--TGA-- 1519
QY 636 QPRAPSRGRLQRAEQVSRALOPALDSYFHPGCTPAPGRGVGPGAG 681
Db 1520 -----AASAPVMLRVLYPKTPTMMVFEPEG 1547

RESULT 9
TLE4_RAT
ID TLE4_RAT STANDARD; PRT; 741 AA.
AC Q07141.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Transducin-like enhancer protein 4 (ESP2 protein).
GN TLE4 OR ESP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=94064640; PubMed=8245004;
RA Schmidt C.J., Sladek T.E.;
RT "A rat homolog of the Drosophila enhancer of split (groucho) locus lacking WD-40 repeats.";
RL J. Biol. Chem. 268:25681-25686(1993).
CC -!- FUNCTION: NUCLEAR EFFECTOR MOLECULE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14463; AAC37640.1;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40. 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40. 6.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Nuclear protein; Phosphorylation; Repeat; WD repeat.
FT DOMAIN 1 106 GLN-RICH (OR Q DOMAIN).
FT DOMAIN 107 173 GLY/PRO-RICH (OR GP DOMAIN).
FT DOMAIN 175 242 CCN DOMAIN.
FT REPEAT 441 472 WD 1.
FT REPEAT 499 529 WD 2.
FT REPEAT 543 573 WD 3.
FT REPEAT 585 615 WD 4.
FT REPEAT 667 697 WD 5.
FT REPEAT 708 738 WD 6.

```

```

FT MOD_RES 218 218 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 233 233 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SQ SEQUENCE 741 AA; 80490 MW; 43F105E568FD87A5 CRC64;

Query Match 2.98; Score 105.5; DB 1; Length 741;
Best Local Similarity 19.3%; Pred. No. 1.5;
Matches 104; Conservative 54; Mismatches 147; Indels 235; Gaps 27;

QY 23 RLVGPDATHSPGLSCRLWSDIILCLPGDIVPAPG--PVLAP--THLOTELVLRCQKET 78
Db 297 RTDAPTFGNSSTPLGR-----PVGKPPGVDPDPLASSURTPMAYPCPPT 340
QY 79 DCDLCLRVAVHLAVHGWEEPEDEKFGGAADSGVEEPRNASLQAQVVLSEFQAYTARCV 138
Db 341 P-----FGIVPHAGMNGELTSP-----GAAYAGL--HNISPMQMSAAA----- 376
QY 139 LLEVQVPAALVQFGSGVGVVYDCFEAALGSEVRISWSTQPRYKELNHTQOLPALP--- 195
Db 377 -----AAAAAAYGRS-----FVVGFDPHHMRVPAIPNL 407
QY 196 -----WLVNSADGDNVHLVNLVNSEEQHGLSLYWNQVGPCKPRHKNLTGP--- 242
Db 408 TGIFGKPAYSFHVSADG-----QMQPVFPF--PDALLGPGIP 443
QY 243 ---QIITLNHTDLVPCV-----CIQVWPLEPDSVYRTNICPFREDPRAHQ 283
Db 444 RHAQINTLNHGEVVCATISNPTRHVYTGKGVKVDITDPGNKSPVSQL--DCLNRD 501
QY 284 NLWQAARL-----RL-----TLQSWLLDAP-----CSLPAERALCHRAFGGDCQPLV 327
Db 502 NYIRSCRLLPDGRITLVGGEASTLSIWDLAAPTPIKAELTSSAPACY----- 549
QY 328 PPLSWENVTVDVNSSEKILQLEC-----LWADSLGPLKDDVLLLETRGPQDNRSLSALE 381
Db 550 -----ALAISSPSKSVCFSCSDGNIAV-----DLHNTLVKQFGHTGASCIDIS 596
QY 382 PSGTSLPSKASTRAARLGEYLLDLSGQCLQLWD--DDLGLAWACPM----- 428
Db 597 NDG-TKLWTGCLDNTVR-----SWDLREGLOLQHDFTSQIFSLGVCYCTGEWLAVGMS 650
QY 429 -----DKY---THKRWALVWLACLLFAAALSLILLKKHAKWLRLLKQDVRSG 475
Db 651 NVEVLHVTKPKDYQLHLHE-----SCVLSLKFAHCGKWF-----VRPG 688

RESULT 10
TLE4_HUMAN
ID TLE4_HUMAN STANDARD; PRT; 766 AA.
AC Q04727; Q9ULF9; Q9NSL3;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transducin-like enhancer protein 4.
GN TLE4 OR KIAA1261.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
RN [2]
RP SEQUENCE OF 44-766 FROM N.A.
RC TISSUE=Brain;
RA Bloembergen H., Boecher M., Brandt P., Meves H.-W., Weil B., Wiemann S.;

```





[illegible]

```

OY 455 LKKHAKGWLRLKODVSGAARGRAALLYSADSGFERLVGALASALCOLPLXAV 514
DB 315 -----EVEGAGAPS-GTWLXSHPTNKRKM--NLGIYNNPLE--V 353
OY 515 DLMSRRELSSAGPVAMFAORROTLQEG-GVYVLLFSPG 552
DB 354 NYWSRPMILNRGSPSGYSDL-----AVVEGOGCLFACLEPCG 388

RESULT 13
M3K6_HUMAN
ID M3K6_HUMAN STANDARD: PRT; 1011 AA.
AC 095382;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 6 (EC 2.7.1.-).
GN MAPK6 OR MAPKK6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99092374; PubMed=9875215;
RA Wang X.S., Diener K., Tan T.-H., Yao Z.;
RT "MAPKK6, a novel mitogen-activated protein kinase kinase, that
RT associates with MAPKK5.";
RL Biochem. Biophys. Res. Commun. 253:33-37(1998).
CC 1- FUNCTION: ACTIVATES THE JNK, BUT NOT ERK OR P38 KINASE PATHWAYS.
CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: AF100318; AAD05304.1; -
DR MIM: 604468; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00220; S_TKC.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 371 629 PROTEIN KINASE.
FT NP_BIND 377 384 ATP (BY SIMILARITY).
FT BINDING 400 400 ATP (BY SIMILARITY).
FT ACT_SITE 494 494 BY SIMILARITY.
SQ SEQUENCE 1011 AA; 112512 MW; F7294BA0587D2EC9 CMC64;

Query Match 2.7%; Score 100; DB 1; Length 1011;
Best Local Similarity 21.2%; Pred. No. 6;
Matches 149; Conservative 91; Mismatches 268; Indels 194; Gaps 38;

```

```

OY 219 -GLSL-----YANOVGPCKPMHKNLIGPOIILNHTDLVPC-----L 256
DB 141 IGMKGLCLLARKGVCKMYTYWD-----GFYLAQIILANDPTQVLAIAEQLYKL 190
OY 257 CIOVWPL-----EPDSVRNTICPFREDP-----RAHONLM-----QAARL 291
DB 191 NAPIVLYSVETFLYLHFRPTPEPPGPPRRAR--FWLHLLDSCQPFRTACQSGQC 248
OY 292 RLFLTQSWLIDAPCSLPAEALCWRAPGDCOP-----LVP-----PLSMENVTVDNS 341
DB 249 LVLVLEMKKVLPLPAKLEVR-----GTDPSVTYTLISLEPEQDIDISSW---TFPVAS 297
OY 342 -----SEKIQLOPCIMADSLGPKADVLLLETRGPQDNLSLAFSGCT-----SLPSKAS 393
DB 298 ICGVASAKRDERCCFLVLPADQVQLCFPSVG--HCQMGCLDAWNTNDSTAPAEBA 355
OY 394 TRARILGYLLDQSGCGLDMDDLGALNACPMQDYIHKRNALVWL--ACLLPFAAL- 450
DB 356 EGAGEMLEFDYETETGRLVGLGKTYGVYA--GDRRTFRVIAIKETIPERDSRFSQPLH 414
OY 451 SLILLKDKHAKGWLRLKODVSGAARGRAALLYSADSGFERLVGALASALCOL- 508
DB 415 EELHLRRLRKKNIVRYL-----GSASQG--GLKIFMEVPG-----GLSLILRSYWG 462
OY 509 PLR---VAVDLMSRRELSSAGPVAMFAORROTLQEGGVYVLLFSP-----G 552
DB 463 PLKNESTISPYTROILOGLYLHDNHVHND-IGDGVNLINTFSGLIKISDFGSKRLA 521
OY 553 AVALCSEWLQGV-----SGPGAGPHDAFASISCVLPDLOGRAP-----GSYVG 599
DB 522 GTPCTETFTFTGLQYMAPEITIDQGRGTG-KAADIWSLACTYIEMATRPPEHLSQQA 580
OY 600 ACPD---RLHPDAVPALFRTVPVFTLPSQLPD-----FLGALQOPRAPRS 642
DB 581 AMFGVMKVKVHPMSSLSAEQAFLRTFEPDPLRASQOTLLGDPLQGGKRSRSS 640
OY 643 GRLOER-AEYVSRLQPALDSTF-----HPPGTP 670
DB 641 PRHAPRPSDAPSASPTPSANSTQSPRCPQAPSQHPPSP 682

RESULT 14
CXAA_MOUSE
ID CXAA_MOUSE STANDARD: PRT; 505 AA.
AC 09WS4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction alpha-10 protein (Connexin 57) (Cx57).
GN GJA10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99262620; PubMed=10329667;
RA Manthey D., Bukuskas F., Kozak G., Willecke K., Lee C.G.;
RT "Molecular cloning and functional expression of the mouse gap junction
RT gene connexin-57 in human HeLa cells.";
RL J. Biol. Chem. 274:14716-14723(1999).
CC 1- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC 1- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC 1- SUBCELLULAR LOCATION: Integral membrane protein.
CC 1- TISSUE SPECIFICITY: LOW LEVELS WERE DETECTED IN SKIN, HEART,
CC KIDNEY, TESTIS, OVARY, INTESTINE. EXPRESSION NOT DETECTED IN
CC BRAIN, SCIATIC NERVE OR LIVER.
CC 1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II).
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: AJ010741; CAB40358.1; -  
DR MGD; MG1:1339969; Gjal0.  
DR InterPro: IPR000500; Connexin.  
DR Pfam: PF00029; connexin; 1.  
DR PRINTS: PR00206; CONNEXIN.  
DR SMART: SM00037; CNX; 1.  
DR PROSITE: PS00407; CONNEXINS\_1; FALSE\_NEG.  
DR PROSITE: PS00408; CONNEXINS\_2; 1.  
KW Gap junction; Transmembrane.  
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 20 40 POTENTIAL.  
FT DOMAIN 41 76 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 77 97 POTENTIAL.  
FT DOMAIN 98 165 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 166 186 POTENTIAL.  
FT DOMAIN 187 209 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 210 230 POTENTIAL.  
FT DOMAIN 231 505 CYTOPLASMIC (POTENTIAL).  
SQ SEQUENCE 505 AA; 57115 MW; 77FB9575FD0A274A CRC64;

Query Match 2.7%; Score 99; DB 1; Length 505;  
Best Local Similarity 19.6%; Pred. No. 2.9;  
Matches 70; Conservative 50; Mismatches 107; Indels 130; Gaps 20;

QY 125 VVLSFOAYPTARCVLLEVOVPAALVQSGSVVYDCFEALGSEVRWISYTPRYEKE 184  
DB 177 ILYGFQHHPIYK-----TQACP-----NSV-DCF-----VSRP----- 205  
QY 185 LNHTQOLPALPWLNVSGADGNVHLVNSEQHFGLSLYWNQVQPPKPRWHKLTGPQI 244  
DB 206 ---TEKTFIMLFHSIA---AISLLNILEIFHGIKIRMLADGKSSGNTENETGPPF 259  
QY 245 IYLNHTDLVPLCIQWVP-----LEPDS-----VRTNICFPREDPRAHONLWQAARLLT 295  
DB 260 HSTNYSGTQOCMVCSLPRISLQANNKQOIVRNI-----PRS-KSMQIPHPRQLE 312  
QY 296 L-----QSWL-----LDAPCSLPAEALCNRAPGDCPQLVPPL----- 330  
DB 313 VDVSCKRQWAKEIESCTQLHVSPC---PHRSARIHQPCQCHSVGFKNMSQSWFG 370  
QY 331 -----SWENVTVVNSSEKQLQLEC-----LWADSLGP--LKD 361  
DB 371 TMTASQHRPSSALETWERSQGPPEASGRSLTDROSHFGSDGSAESGVWTDRLGPGSRKA 430  
QY 362 DVL--LLETRGPQ-----DNRSICALPEPSC-----TSLPSKASTRAARLGE 401  
DB 431 SFLSRLMSEKQRHSDSGSSRL-----NSSCLDFSHGNSPSPPLSATGHRASMSVK 483

RESULT 15  
W70T\_MOUSE  
ID W70T\_MOUSE STANDARD; PRT; 922 AA.  
AC Q9D2V7;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 70 kDa WD-repeat tumor rejection antigen homolog.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombers P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -I- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
CC -I- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AK018739; BAB31380.1; -  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 6.  
DR PRINTS: PR00320; GPROTEINRPT.  
DR PRODOM: PD000018; WD40; 2.  
DR SMART: SM00320; WD40; 6.  
DR PROSITE: PS00678; WD\_REPEATS\_1; 1.  
DR PROSITE: PS00882; WD\_REPEATS\_2; 4.  
DR PROSITE: PS0294; WD\_REPEATS\_REGION; 2.  
KW Repeat; WD repeat.  
FT REPEAT 75 115 WD 1.  
FT REPEAT 124 163 WD 2.  
FT REPEAT 166 205 WD 3.  
FT REPEAT 209 253 WD 4.  
FT REPEAT 293 331 WD 5.  
FT REPEAT 331 369 WD 6.  
FT REPEAT 369 407 WD 7.  
FT REPEAT 407 445 WD 8.  
SQ SEQUENCE 922 AA; 100860 MW; C8E4BC1EE9CFE242 CRC64;

Query Match 2.7%; Score 99; DB 1; Length 922;  
Best Local Similarity 20.3%; Pred. No. 6.4;  
Matches 99; Conservative 48; Mismatches 120; Indels 220; Gaps 23;

QY 12 LGRSPVLSRLRYCPQ-DATCSPGLSCRLWSDILCLPGDIVPAPGPVLAPTHQLTEL 70  
DB 451 IGTSPSRQSLSLGLPGSCFKFRHTQGSLLHR---DSHTNLKLNLTTPG----- 496  
QY 71 VLRCOKETD--CDLCLRVAVHLVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVL 128  
DB 497 -----ESDGFANRLRVAVPLSSG----- 516  
QY 129 FOAYPTARCVLLEVOVPAAL-----VOFGSGSVVYDCFE-----AALGSEVRWIS 175  
DB 517 -----GEVAVLELQKPGRLPDTALPTLQNGTAVMDLVMDPDHRLAVAGEDARILWR 570  
QY 176 YTPRYEKE-----NHTQOLPALPWLNVSD---GDNVHLVNV----- 212  
DB 571 VPPGLENVLTPTPETVLTGHTKEIYSLFRHPLAADVLASSSYDLTVRWDLTGAEUKL 630  
QY 213 --SEQHFGLSLYWNQVQPPKPRWHKLTGPQIITLNLHTDLVPLCLCIQVWPLEPDSVRT 270

```
Db 631 QGHQDQIF--SLAWSP-----DGKOLATV-----CKDGHVRVYEP---RS 665
Qy 271 NICPFREDPRAHQNLWQAARLLTLQSWLLDAPCSLPAEAAALCWAPGGDPCQPLVPPL 330
Db 666 SPLPLQEGGPEGG--RGARI-----VWVCDGGCLLV----- 696
Qy 331 SWENVTVDNSEKIQEQECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPS 390
Db 697 -----GFDSREROLOL---YIADAL-----AQGPS---ALLGLDVAPOSTLLPS 734
Qy 391 -----KASTRAARLGEYL-----LQDLQ 408
Db 735 YDPDTGLVLLTGKGTTRVF-LYEVLPFAEPFFLECNSTSPDPHKGFVLLPKTECDIQDVE 793
Qy 409 SGQCLQL 415
Db 794 FARCLRL 800
```

Search completed: May 19, 2002, 17:01:06  
Job time: 10247 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:11:33 ; Search time 63.07 Seconds  
(without alignments)  
1048.192 Million cell updates/sec

Title: US-09-608-918-10  
Perfect score: 3692  
Sequence: 1 MPVFWFLSLALGRSPVLS.....TPAPRGVGPAGAGDGT 688  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	126	3.4	3707	2 S18252	heparan sulfate pr
2	125.5	3.4	1366	2 T35985	probable large pr
3	124.5	3.4	761	2 S20458	pqgf protein - kle
4	120.5	3.3	478	2 D75564	hypothetical prote
5	115.5	3.1	698	2 A82593	hypothetical prote
6	114.5	3.1	1711	1 A55148	protein-tyrosine-p
7	112.5	3.0	1193	2 F83264	hypothetical prote
8	111	3.0	1118	1 A49724	protein-tyrosine-p
9	109.5	3.0	2055	2 T00093	hypothetical prote
10	108.5	2.9	753	1 WZBE8	gene 56 protein -
11	107.5	2.9	433	2 B75376	sensor histidine k
12	107	2.9	825	1 A60386	interleukin-4 rece
13	106.5	2.9	274	2 A83583	probable biotin sy
14	105.5	2.9	679	2 B75262	conserved hypothet
15	105.5	2.9	741	2 B49555	enhancer of split
16	105	2.8	724	2 T47149	hypothetical prote
17	105	2.8	3670	2 T36249	CDA peptidase synthe
18	104.5	2.8	238	2 T070841	hypothetical prote
19	103.5	2.8	1085	2 T03531	cobn protein homol
20	102	2.8	2403	2 A59386	sanko - human
21	102	2.8	2629	2 T30987	telomerase-associat
22	101.5	2.7	418	2 JC7588	exo-alpha-stalidas
23	101.5	2.7	859	2 T35785	probable beta-gluc
24	101.5	2.7	1737	2 T00209	MEGF8 protein - hu
25	101	2.7	427	2 E87669	conserved hypothet
26	100.5	2.7	745	2 H85840	hypothetical prote
27	100	2.7	719	2 T35189	probable ATP-depen
28	100	2.7	846	2 T27282	hypothetical prote
29	99.5	2.7	745	2 C99995	hypothetical prote

ALIGNMENTS

RESULT 1

S18252  
heparan sulfate proteoglycan - mouse  
N:Alternate names: perlecan  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S18252; A31917; B31917; S66460  
R:Noonan, D.M.; Fullie, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.;  
J. Biol. Chem. 266, 22939-22947, 1991  
A:Title: The complete sequence of perlecan, a basement membrane heparan sulfate prote  
adhesion molecule.  
A:Reference number: S18252; MUID:92078153  
A:Accession: S18252  
A:Molecule type: mRNA  
A:Residues: 1-3707 <NO>  
A:Cross-references: EMBL:W77174; NID:g200295; PIDN:AAA39911.1; PID:g200296  
R:Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.; H  
J. Biol. Chem. 263, 16379-16387, 1988  
A:Title: Identification of cDNA clones encoding different domains of the basement mem  
A:Reference number: A92680; MUID:89034110  
A:Accession: A31917  
A:Molecule type: mRNA  
A:Residues: 940-1601 <NO2>  
A:Cross-references: GB:J04054; NID:g200252; PIDN:AAA39899.1; PID:g200253  
A:Accession: B31917  
A:Molecule type: mRNA  
A:Residues: 1870-2600 <NO3>  
A:Cross-references: GB:J04055; NID:g200300; PIDN:AAA39912.1; PID:g200301  
R:Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpi, R.  
Eur. J. Biochem. 231, 551-556, 1995  
A:Title: Structural properties of recombinant domain III-3 of perlecan containing a g  
A:Reference number: S66460; MUID:95377282  
A:Accession: S66460  
A:Molecule type: protein  
A:Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>  
A:Superfamily: LDL receptor ligand-binding repeat homology; EGF homology; laminin G r  
C:Keywords: glycoprotein  
F:199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F:368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F:764-811/Domain: laminin-type EGF-like homology <LEG>  
F:1159-1206/Domain: laminin-type EGF-like homology <LEG7>  
F:1563-1610/Domain: laminin-type EGF-like homology <EG7>  
F:1613-1668/Domain: laminin-type EGF-like homology <LEG8>  
F:3163-3198/Domain: EGF homology <EGF>  
F:3270-3423/Domain: laminin G repeat homology <LEG2>  
F:3464-3492/Domain: EGF homology <EGF7>  
F:1256,1891,2336,2394,2427/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 3.4%; Score 126; DB 2; Length 3707;



[illegible]

RESULT 4  
D75564  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: D75564  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shan, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Uitterlacker, T.; Zaleski, C.; Mc  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: D75564  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-478 <WHI>  
A:Cross-references: GB:AE001869; GB:AE000513; NID:g6457721; PIDN:AAF09658.1; PID:g6457721  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0061  
A:Map position: 1

	Query Match	3.3%	Score 120.5	DB 2	Length 478
	Best Local Similarity	24.3%	Pred. No. 0.21		
	Matches	99	Conservative	39	Mismatches 146
					Indels 121
					Gaps 17
QY	276	REDPRAHONLWQAAARLRLTLTQSMIL-----DAPCSLPAAEALCWPAGDPPQPLV	327		
Db	110	RADPREFRLPREEAAO-----RIRRMRLRLARAPGOLPGTVPAHSVLY-----AOEASALS	161		
QY	328	PLSLRENTVYDVNDSSEKLOLOECL---NADSLGRLKDDVLLTFTRGPQONRSLCALPESG	384		
Db	162	APAAGGGLLHLHGPRDLAHWQGLRLSNW-SLGPMMQOLLYODETSOPVSGGQALLRP--	218		
QY	385	CTSLPSKASTRAARLGEVLLDLDLQSGOCLQMDDDLIGALMACPMYKIHKRMALVWLACT	444		
Db	219	--DLPAERQQLFLVGGQQLQVLFSG-----	242		
QY	445	LFAAALSLILLKKKHAKCWMRLLLKQDVRSGAARGRA-ALLLYSADDSGFERLYGATAS	503		
Db	243	-----DYVLLRRRDAOTAOTRLAR-----LAAIGRCACALLLPREQGRILRLAATLAR	290		

```

Oy 504 ALCCPLPLVAVDLMSRRSLSSQGPAMFMAHRRQTLDDGGVVLLFSPGAV-----554
Db 231 RLKGDPPR-ADDPASGGVAFKQTPRECCLAARR-SLQRIEPLTLRFSPQAQYHNAEQAOQA 348
Oy 555 -----ALCSEMLQDVGSP--GAHGPHDAFRASLSCVL-----PDFL 589
Db 349 LLLHLPPIHAAOYQQAHEHAASAEPLKRAATDPLVGRACGPPVLPATGHPIVFPPLGFDPL 408
Oy 590 QGRAGSGYVGACFDRL--LHDDAVPALERTVPVETLSQLPDFIGAL 634
Db 409 EYRLPLG-----DRLLTLRQD-----YRALVAVLPEQQAAYVYGDLL 443

```

RESULT 5  
A82593  
A82593  
Hypothetical protein XF2169 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: A82593  
R:annotatus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: A82593  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-698 <SIN>  
A:Cross-references: GB:AE004030; GB:AE003849; NID:g9107292; PIDN:AAF84968.1; GSPDB:GNF5  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Briones, M.R.S.; Bueno, M.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carren  
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; La  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, G.  
A:Authors: Martins, E.M.F.; Matsumura, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
Rodrigues, V.; Rosa, A.U. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva  
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2169

Query Match	3.1%	Score 115.5	DB 2	Length 698
Best Local Similarity	21.8%	Pred. No. 0.84		
Matches 160	Conservative	90	Mismatches 234	Indels 251
			Gaps	43

  

QY	61	LAP----	THLQELVLRCKE--TD	CDLCGRVAVHLAVHGWHEDEE	KFG-----	G	107
DB	69	LAPERPELHLQGLILL	LNQHDLSGDAALSR	TRL-----	DP	QFNYYVQAHIA	118
QY	108	AADSCVEERBNASLO	QAVVLSFOAYP----	TARCVLLEQVAPALVQ	FGQSV-----	G	156
DB	119	IARGALDEAEQSL	TRATRIIS--DH	QQLAVNATIALRGQVQ	DQALMTLSQAVELL	PEEP	176
QY	157	SVVYVCFEAAAGSE	VRKMSYTOPRY----	EK-----	ELN--HTQDLRALPYL	NSADGN	205
DB	177	AVLF-----	SLG-----	FAYLQKGHAFAERA	FORVTEINP	HYPAVRLAQLAORQ	R- 225
QY	206	VHLVAVNSEEHQF	GLVYVNOVGPRK--	PRMYKVLTPROJIT	LNHTDPLVPSLC	IOV--	260
DB	226	LDDARIRIEE--	ASLS-----	PEGDPRALNR--	LTGEFEELAGHP	RALTYHLRQVLA	273
QY	261	WPLEPDSRVNTICE	PRE-----	DPRAHQIMQAA	RURL-----		293
DB	274	TWPEEPRTLEALT	AMKQKQMDDDAR	ITILDAALDIKPR	NH--DLN--LAR	LAVAPVSGDEAR	331
QY	294	LTLOSMLDAPCS	LPAEAAALCWAP	RAGBQDQCPRLV	PLPSMENVTYDV	VNSSEKLODECTMA	353



```
Db 332 IVIERKLSAMPHELPALETLM----- 352
Qy 354 DSLGLPKDDVLLLETRGPQDNRSLSALEPSGCTSLPSKASTRAARLGEVLLQ--DLQSGQ 411
Db 353 -SLHDIONNPEAAETVA---RQIVALEPGRLSG-----EQRIVEALLQRPPTAVA 399
Qy 412 CLQWDDDLGALWACPMCKYIKHRWAL-VWLACL-----LFAAASLIILLKKDHAKGWL 465
Db 400 CLQOLIESVPE-----HERTILRWLGLVQDRAGQFEAALATWLQFOAEQAKYRL 449
Qy 466 RLLKQ-----DVRSCAAARGRAALLYSADDSGFRRLYGALA--SALC----- 506
Db 450 PLPPTWNSKOMPDLAIPAELARPLLIWGPVSHVETVMMMAWASMLCADRYSET 509
Qy 507 -----QLPLRVAVDLMSRRELSAOGP-----VAWFH-----AQR 536
Db 510 PPADPLORYETVSELTSGTLTPQALIDAW-RQOLPARGIEDGNVIDWLWMDNSLITALR 568
Qy 537 QTLQEGGVVVLLFSPGAVALCEWLQDGVSGAGHPHDAFRASLSCLVDFLQGRAPGS 596
Db 569 PHLPEGRLLIIVLRDPRMLL--DWIAYGSPILPA---LDSLQQAANW-LGDIL-----N 616
Qy 597 YVGACFDRLHP-----DAVPALFRTV-PVFTLPSQLPDFLQALQQPRAPRSGL 645
Db 617 QIAALHELDLYPHHLIRLDGIEDNPQALATTLEDIFGSPFPIPP---SLEAPRLP-AGRW 672
Qy 646 QERAEQVSRALQPAL 660
Db 673 RDYREVLSAFDAVL 687

RESULT 6
A55148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type OST precursor - rat
N:Alternate names: OST-PRP; osteotesticular protein-tyrosine-phosphatase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55148
J:Mauro, L.J.; Olmsted, E.A.; Skrobacz, B.M.; Mourey, R.J.; Davis, A.R.; Dixon, J.E.
R: Biol. Chem. 269, 30659-30667, 1994
A:Title: Identification of a hormonally regulated protein tyrosine phosphatase associated
A:Reference number: A55148; MUID:95074080
A:Accession: A55148
A:Molecule type: mRNA
A:Status: not compared with conceptual translation
A:Residues: 1-1711 <MAU>
A:Cross-references: GB:L36884
C:Comment: The sequence contains ten fibronectin type III repeats and two protein-tyrosine
C:Superfamily: protein-tyrosine-phosphatase, receptor type OST; fibronectin type III rep
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1711/Product: protein-tyrosine-phosphatase, receptor type OST #status predicted <MA
F:1174-1398/Domain: protein-tyrosine-phosphatase, homology <PTPI>
F:1350/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1356/Binding site: substrate phosphate (Arg) #status predicted

Query Match 3.1%; Score 114.5; DB 1; Length 1711;
Best Local Similarity 19.8%; Pred. NO. 3.1;
Matches 158; Conservative 73; Mismatches 270; Indels 295; Gaps 39;

Qy 5 WFLLSLALG-RSPVLSLERLVGPQDATHCSPLGSLRWLSDILCLPGDIVAPGVLAP 63
Db 322 WASHKAGLGARDGVLV---KLSGPMESTSLGPESCN-----AVFPGP-LPP 364
Qy 64 THLQTELVRQCKETDCDCLRLRVAVHLAVHGWEEPEDEKFGGAADSGVEE----PRNA 119
Db 365 GHYTLQL-----KVLAGPYDAWVEGSTWLAESEA 392
Qy 120 SLQAOVV-----LSQAYPTARCVLLEVOVPAALVOFGSGVS--VYVDFCEAALGS 169
Db 393 ALPREVPGARLWLDGLEASKQPGRRALLYSDAPGSLGNISVPSGATHVIFGCLVPGAHY 452
```

```
Qy 170 EVRIWSTQPRYERKELNHTQOLP--ALPWLNVSADGDNVHLVNLVNSEQHFGLSLYWNQV 227
Db 453 RVDIASSTGDISQSISGYTSPPLPQSLSEVISRSPSD-----LTIWGA 497
Qy 228 QGP---PKPRHKNLT-----GPQIITLNHTDLP--CLCIQVWP---LEPDSV 268
Db 498 PGQLEGYKVTWHODGQSQRSPGDLVDLGDPTLSLTLSLVPGSSYTVSAWAGNLGSDSQ 557
Qy 269 RTN1CP-----FREDPRAHONLW-----CAARLRLTLQ----- 297
Db 558 KIHSCTRPAPPTNLSLGFHQPAALKASWYHPPGGRDAFHRLRLRLTLTLESEKVLDPRE 617
Qy 298 ----SWL-LDAPCSLPAEALCW-----RAPGGDPCQLVPPLSWENTVVDVNSSEK 344
Db 618 AQNFSWAQLTAGCEQVQLSTLWGSERSSSANATGWP--PSAPL-----VNVTSAP 669
Qy 345 LQLOECUWADSLGPKDDVLL-----ETR-----GP-ODNRSLSCALEPS-----GCTSLP 389
Db 670 TOLQ-VSWAHVPGGRSRYQVTLYQESTRTATSIMGPKEDGTSLGLTGTGYKKEVISWA 728
Qy 390 SKASTRAARLGEY-----LLQDLQSGCQLQWDDDLGALWACPMCKYIKHRWALVWL 441
Db 729 GPLYTAANVSAMTYPIPLINELIVSMQAGS-----AVVNL 763
Qy 442 ACLLFAAALSLLILLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSADDSGFRRLVGLAL 501
Db 764 A-----W-----PSGPLGGGACHAQLSDAGHLSWEQ----- 789
Qy 502 ASALCOLPLRVAVDLMSRRELSAOGPVAMFHAQRQRTLOEG-GVVVLLFSPGAVALCEW 560
Db 790 -----PLKLGOELFMLRLDLPCHTISMSVRCRAGPLQASTHLVWLVSVEPGPV----- 836
Qy 561 LDGVSFGAGHPHDAFRASLSCLVLPD-----FLOGRAPGS-----YVGACFDRL 606
Db 837 -EDVLCHP-----EATYALALNMTMPAGDNDVCLVVERLVPPGGGTHFVQVNTSGDALL 889
Qy 607 HPDAVPALEFVTPVFTLPSQLPDFLQALQQPRAPRSRQLOERAFQVSRALQPALDSYFHP 666
Db 890 LPNLMPTTSYRLSLTVL-----GRNSRWSRAVSLVCSLSAEAW-HP 929
Qy 667 PGTPAP-----GRGVG 677
Db 930 PELAEPPQVELGTGMG 945

RESULT 7
F83264
hypothetical protein PA3063 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83264
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: F83264
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1193 <STO>
A:Cross-references: GB:AE004730; GB:AE004091; NID:g9949154; PIDN:AG06451.1; GSPDB:G
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3063

Query Match 3.0%; Score 112.5; DB 2; Length 1193;
Best Local Similarity 23.0%; Pred. No. 2.8;
Matches 155; Conservative 60; Mismatches 276; Indels 183; Gaps 38;

Qy 81 DCLRLVAVHLAVHGHWE-----EPEDEKFGGAADSGV---EEPNLSLQAOVV--- 126
```

```

Db 293 DLRLNMAAG-RVERAWEVGOELLALRPEDRTLLADLARLGWTGNGPRALGEWQALLGA 351
Qy 127 -----LSFOAYPTARCVLLEVVQVPAALVQGSQVSVVYDCFEALGSEVRIW- 174
Db 352 DDPALREHAWRLSLQMFDFDSATIELLAPIGAQRQMTDEELDALVYSHETRTGTPGEAWL 411
Qy 175 -SYTOPRYEKE-----LNHTQOLPALP--WLNVSADGDNVHLVNVSEHQECLS 221
Db 412 RGYVO-RYPQRLAWORLQOILEHTQOLQOEBETGVWARMAR-----HPLSVKERMQNA-E 464
Qy 222 LYWNQVQGPCKPRWKNLKGQPIITLHNTDLVPCLCIQVWPLEPDSVVRTNICPPREDPRA 281
Db 465 THWNLFDP--PROAW-KVLAVDTRAIREPEFWRLRAALAWALEQD-----DAR- 510
Qy 282 HONLWQAARLLTLTQSWLLDAPCSLPAAALCWAPAGPCQPLVPP-L-SWENVTVDVN 340
Db 511 -----AAYERMLA-----LDIRLSRDEDLIALYRDSNPKQALQVLIGSWQRSDPRR 559
Qy 341 SSEKLQLOECL--WADSLGLKDDVLLLETRGPQDNRSICALPSGCTSLPSKAST----- 394
Db 560 LASALQLAENLHWP-----ALKS--LLAEABG-----LPEAQGSPYYW 596
Qy 395 -RAARLGEVLLDQSGQCLOLWDDDLGALWACPMCKYIHKRWALVW-----LAC 443
Db 597 VARARLAE--QEGHGDVAERLYRE--ALVRPPEGNLVRE--LLWFYIDRGRDRLSLAP 648
Qy 444 LL-----FAAALSLILLKKHAKGWRL-LKQDVRSGAAAGRAALLL 486
Db 649 LLAQHWGLALRDLTLWLPFASA--SELLERNQALAWFLYLSKSNPNDLWVQAYADAL- 705
Qy 487 YSADSGFE-----RLVGALASALCQLPLRVAVDLWS-----RRELSAQGPVAMFHAQRQTL 539
Db 706 ---DASGYQDKALRLRLRLRLRDREAVRATPDSFATYLRLLAVAQGPLLAQGEARRAWN 762
Qy 540 QEGGVVWLLF-----SPGAVALCSEWLDGVSGPAH-GPHDAFRASLSCLVPLDFLO 590
Db 763 GPAMLLQWLFEFLDQLAATNOEPLKDNWLA-WARGRLKIGRNEEQAAALRS-----Q 815
Qy 591 GRAPGSYVGCACFDRLHPDAPALFTVPVFTL---PSQLPDFLQALQOPRAPSRLQE 647
Db 816 NR-----AALQRLERELDELPQAQRVEALVRLGHGGEALGAL-----GDGHSRD 862
Qy 648 RAEQVSRALQPALD 661
Db 863 NREQLRRQAEEILE 876

```

```

RESULT 8
A49724
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type H precursor - human
N:Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated type 1; SAP-1
C:Species: Homo sapiens (man)
C>Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 16-Jun-2000
C:Accession: A49724
R:Matczak, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, K.; Horita, K.
J. Biol. Chem. 269, 2075-2081, 1994
A:Title: Molecular cloning of a human transmembrane-type protein tyrosine phosphatase an
A:Reference number: A49724; MUID: 94124561
A:Accession: A49724
A:Molecule type: mRNA
A:Residues: 1-1118 <MATO>
A:Cross-references: GB:D15049; NID:q475003; PIDN:BAA03645.1; PID:q475004
C:Genetics:
A:Gene: GDB:PTPRH; SAP-1
A:Cross-references: GDB:305504
A:Map position: 19q13.4-19q13.4
A:Note: highly expressed in colon and pancreatic cancer cells but not in the normal cell
C:Superfamily: protein-tyrosine-phosphatase, receptor type H; fibronectin type III repea
C:Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric monoox
F:1-27/Domain: signal sequence #status predicted <SIG>
F:27-110/Domain: fibronectin type III repeat homology <3FNA>
F:28-1118/Product: protein-tyrosine-phosphatase, receptor type H #status predicted <MAT>

```

```

F:28-761/Domain: extracellular #status predicted <EXT>
F:116-199/Domain: fibronectin type III repeat homology <3FNB>
F:205-289/Domain: fibronectin type III repeat homology <3FNC>
F:296-379/Domain: fibronectin type III repeat homology <3FND>
F:385-468/Domain: fibronectin type III repeat homology <3FNE>
F:474-558/Domain: fibronectin type III repeat homology <3FNF>
F:564-658/Domain: fibronectin type III repeat homology <3FNG>
F:667-737/Domain: fibronectin type III repeat homology <3FNH>
F:762-778/Domain: transmembrane #status predicted <TMN>
F:779-1118/Domain: intracellular #status predicted <INT>
F:846-1070/Domain: protein-tyrosine-phosphatase homology <PTPI>
F:35,78,83,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,55
F:1022/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1028/Binding site: substrate phosphate (Arg) #status predicted

```

```

Query Match 3.0%; Score 111; DB 1; Length 1118;
Best Local Similarity 18.4%; Pred. No. 3.4;
Matches 126; Conservative 65; Mismatches 200; Indels 294; Gaps 31;

```

```

Qy 56 APGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEK-----FGGA 109
Db 387 APNPV-RNLHMET-----QTNSSIALC-----WEVPDGPYQDVTYVWGYTG 427
Qy 110 DSGVEEPRNASLOAQVLSFQAYPTARCVLLEVVQVPAALVQGSQVSVVYDCFEALG 169
Db 428 DGGGTETRT-----TNTSVTAERLEPGTLYTFS-----VMAEKNRGARG 467
Qy 170 EVRIWSYTOPRYEKELNHTQOLPALPWLNVNSADGNVHLVNVSEHQEGLSLYWNVOG 229
Db 468 QNVVISITVP-----NAVTSLSKODWNTST-----IALRTAPOG 502
Qy 230 PKPRW-----HKNLTGPQIITLHNTDLVPCLCIQVWPLEPDSVVRTNICPFREDPRAH 282
Db 503 PQGSSYSVSWVWREGMTDPRQTSGTD-----ITLKELEAGSL-----YH 544
Qy 283 QNLWQAARLLTLTQSWLLDAPCSLPAAEA-----ICWRAPG----- 319
Db 545 LTVW-AERNEVRGVNSTLTAA--TAPNEVTDLQNETQTKNSVNLWMLWAKPGDPHSQLYVW 601
Qy 320 -----GDCQPLVPLSWENVTVDVNSS-EKLOLQE-----CLWADSLGLKDDVLL 365
Db 602 VOMASKGHPRRGQDPQANWVWNTSRTNETWYKVEALEPGTLYNFTWAE-----RNDV-- 654
Qy 366 LETRGPQDNRSICALPSGCTSLPSKASTPRAARLGEVLLDQSGQCLOLWDDDLGALWA 425
Db 655 -----ASSTQSLCASTYPTVTITSCVSTAGY-----GVNLIWS 689
Qy 426 CPMCKY-----IHKRW----- 436
Db 690 CPQGGYEAFELEVGQSGQDRSSCGEAVSLGLGPARSYPATITIMDKMKWSHVVC 749
Qy 437 -----ALVWLACLLFAAALLSLILL-----KKDHAKGWLRL----- 468
Db 750 HTESAGVITAGAVGILLFLILVGLLIFFLKRNRKKKQKQPELRDLVFSPPGDIAPDAFAD 809
Qy 469 -----KQDVRSGAA-----ARGRAALLYSADDSGFELVGLGALASALCQLPLRVAV 514
Db 810 HVRKNERDSNCGFADEYQQLSLVGHSSQSQMVASSENNAKNRYRNVLPYDWSRVPKPKTH 869
Qy 515 D-----LWSRRE--LSAQGPVAMFHAQRQTL-----QEGGVVWLL-- 548
Db 870 EEPGSDYINASFMPGLWSPQEFIAQGPLP-----QTVGDFWRLVWFEQSHSTLMLTN 922
Qy 549 -FSPGCAVALCSEWLDGVSGPGAHG 572
Db 923 CMEAGRVCHEWYPLD--SQPCTHG 945

```

```

RESULT 9
T00093
hypothetical protein KIAA0467 -- human (fragment)
C:Species: Homo sapiens (man)

```

```

RESULT      10
WZBER8
      gene 56 protein - equine herpesvirus 1 (strain Ab4p)
C.Species: equine herpesvirus 1
C.Name: host Equus caballus (domestic horse)
C.Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C.Accession: B36801
R.Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A.Description: The DNA sequence of equine herpesvirus-1.
A.Reference number: A36805
A.Accession: B36801
A.Molecule type: DNA
A.Residues: 1-753 <TEL>

```

K:reilford, E.A.R.; Watson, M.S.; McBride, K.; Davidson, A.G.  
Virology 189, 304-316, 1992  
A:Title: The DNA sequence of equine herpesvirus-1.  
A:Reference number: A41831; MUID:92295566  
A:Contents: annotation; possible protein-coding frames  
A>Note: neither amino acid nor nucleotide sequence is given  
C:Genetics:  
A:Gene: 56  
C:Superfamily: varicella-zoster virus gene 54 protein

Query Match 2.9%; Score 108.5; DB 1; Length 753;  
Best Local Similarity 22.1%; Pred. No. 3.2;  
Matches 100; Conservative 46; Mismatches 164; Indels 143; Gaps 22;

QY 295 TLQSWLLDAPCSLPAEALCWRA---PGGD-----PCQPIVP-----P 329  
||| ||| :  
Db 156 TLQSLLDVRSITA---CFSASEPDGTASFAYKVIDWIAQLGLIQLQRLKRAPGATVHP 211  
||| ||| :  
QY 330 LSWENVTVDNSSKQLQLECLWADSLGLKDDVLLLTETRPQDNRSI----- 377  
||| :  
Db 212 KLWRKLPDVPs-----LESCV-----DERDLAGKLYVANSLLR 245  
||| :  
QY 378 ----CALPSPGCT-SLPKASTRAARLGEYLIQDL-----QSGCGLQIWD---DDLQ 421  
||| :  
Db 246 EGLEAVVELARCTASVAMDYDRVNIFFHYTRREVVAIDSTGKRGECVLWQPIWKDGS 305  
||| :  
QY 422 ALWACPMDKY-----IHKRALVWLACLLFAAALSLLILLKNDHAKGWLRLKQDVRS 475  
||| :  
Db 306 VLFDSPLOIRGCEVCNCHALREHAKLCQLLNTVPVKILVGRKKDEAQ-----PG 355  
||| :  
QY 476 AARGRAALLLYSADDSGFERLVGALASALCOLPLRVAVDLWRSRELSAQGPVAFHAQR 535  
||| :  
Db 356 WASKAVDKLM-----GEGEELHSSAASRL----VKLVNKMWRHIGDITETVRSYLN 406  
||| :  
QY 536 ROTLQEGGVVLLFPSPGVALCSEWLQDGVSGPAHGP-HDAFASLSCVLPDFLOGRAP 594  
||| :  
Db 407 TSTNLLSQAQVDTSLPG-----FGQSGKTKQGNMPVQEAFTSYVINGMLEG--- 456  
||| :  
QY 595 GSYGACGAFDRLLHPDAVPALEFTVPVFTPLSQLPDFLGAQQPAPRSGRLQERAEQVSR 654  
||| :  
Db 457 --YVN-----NLFKTIEDLRGTNS-----GLLDQLR-DRESEITHLREQLLR 495  
||| :  
QY 655 ALQAPALDSYFHPGTPA--PGRGVGPAGPGAG 685  
||| :  
Db 496 VSQAADGSTQPGSSAALPGSAKSGAG-GLG 527  
||| :

RESULT 11  
B75376  
sensor histidine kinase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75376  
R:White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson  
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski

RESULT 11  
B75376  
sensor histidine kinase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75376  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75376  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-433 <WHI>  
A:Cross-references: GB:AE002004; GB:AE000513; NID:g6459366; PIDN:AAF11166.1; PID:g645937  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1606  
A:Map position: 1  
C:Superfamily: hypothetical protein HII707; sensor histidine kinase homology

Query Match 2.9%; Score 107.5; DB 2; Length 433;  
Best Local Similarity 22.8%; Pred. No. 1.9;  
Matches 106; Conservative 57; Mismatches 154; Indels 147; Gaps 24;

Qy 289 ARRLRLTQSWLLDAPCSLPAEALCWAPGGD-----PCQPLVPLSWENVTVDVNS 341  
Db 5 ARRLTLTVTVLLSFSFAELAGWLLKAEGLSLTRQLNAQSQALLAVAQSRSTPLDATA 64  
Qy 342 SEKL-----OLOECLWADSL-GPLKDDVLLLETRGPDNRSLCALPSPGCTSL 388  
Db 65 ADILEKGTAAAGRIYQGTLRWGGAGPPVLDAAFLQTSEPR-----AVERVGDILV 118  
Qy 389 PSKASTRAARLGEYLQDQSGCQLQWDDDLGALWACPMKDYIHKRWALVWLACILFAA 448  
Db 119 ASR-----ROGEYV---IQVGNLE-----PLE-HLLRRYALLSGLSLL-- 153  
Qy 449 ALSLILLLKDHAKGWL---RLKQDVRSNAARGRAALL-----LYSADSGFERLV 498  
Db 154 ALSLL-----AGWVAQAKVRRTLRLESVGRVQQLDSPDIPALEQPD-----V 199  
Qy 499 GALASALCOLPLRVAVDLWSRRLSAGQPVAFHQR-ROTLQEGGVVVLFSFGAVALC 557  
Db 200 GALARAL-----QSSLDALRAERETFLASASHLRTPTVTAMLA 240  
Qy 558 SEWLQDQVSGGAGHPHDAFRASLSCLVPLDFLOGRAGSYVVGACFDRLHHPDAPAL--- 614  
Db 241 D-VQHTLSRP---RPPDELRAAL-----MRAEQTASRLRLTLGNLMTLWTWQSLPEP 288  
Qy 615 ----FRTVPVFTLPSQLPD-----FLGALQOPRAPSRRGRLOERAEQ-VSRALQ 657  
Db 289 RAARQOLDLHLAGEAVDRLQPLSLGRVWELWLDGQEPVSGEAGLLSSVAENLIGNAIK 348  
Qy 658 PALDSYFHPG-----TPAPCRG-----VGPAGAGGAGDG 687  
Db 349 -----FTPAGGEVQVRVTLPLNGGAELVVQDTGPGFPPTLTG 386

RESULT 12  
A60386  
Interleukin-4 receptor precursor - human  
A:Alternate names: IL-4 receptor  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 22-Jun-1999  
C:Accession: A60386; A47603  
R:Galizzi, J.P.; Zuber, C.E.; Harada, N.; Gorman, D.M.; Djossou, O.; Kastelein, R.; Banc  
Int. Immunol. 2, 669-675, 1990  
A:Title: Molecular cloning of a cDNA encoding the human interleukin 4 receptor.  
A:Reference number: A60386; MUID:91120547  
A:Accession: A60386  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-825 <GAL>  
R:Idzerda, R.L.; March, C.J.; Mosley, B.; Lyman, S.D.; Vanden Bos, T.; Gimpel, S.D.; Dir  
J. Exp. Med. 171, 861-873, 1990  
A:Title: Human interleukin 4 receptor confers biological responsiveness and defines a nc  
A:Reference number: A47603; MUID:90171849  
A:Accession: A47603  
A:Molecule type: mRNA

A:Residues: 1-74, 'I', 76-825 <ID2>  
A:Cross-references: GB:X52425; NID:g33833; PIDN:CAA36672.1; PID:g33834  
C:Genetics:  
A:Gene: GDB:IL4R  
A:Cross-references: GDB:118823; OMIM:147781  
A:Map position: 16p12.1-16p12.1  
C:Superfamily: interleukin-4 receptor; cytokine receptor homology  
C:Keywords: cytokine receptor; glycoprotein; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-825/Product: interleukin-4 receptor #status predicted <MAT>  
F:26-232/Domain: extracellular #status predicted <EXT>  
F:34-218/Domain: cytokine receptor homology <CRS>  
F:233-256/Domain: transmembrane #status predicted <TM>  
F:257-825/Domain: intracellular #status predicted <INT>  
F:53,98,128,134,176,209/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.9%; Score 107; DB 1; Length 825;  
Best Local Similarity 19.0%; Pred. No. 4.6;  
Matches 154; Conservative 75; Mismatches 270; Indels 312; Gaps 38;

Qy 41 LWDSILCLPGDIVP-----APGVLAPTHLQTELVLRCQKQETDCDCLRVAVHLAVH 93  
Db 103 LWAGQQLLWKGSKFSEHVKPRAPGNLTVHTNVSDFLLLTWSPYPPDNLNHLTYAVN 162  
Qy 94 GHNEEPDEEPPGGAADSGVEEPRNASLOAQVVLFSQAYPTAKVLE--VOVPAALVOF 151  
Db 163 -IWSENDPAD-----FRIY---NVITYLEPSLRIAASTLKS 193  
Qy 152 GOSVGSVVYDCFEAALGSEVRIMSYTPRYEKELNHTQOLPALPWLWNSADGNDVHLVLN 211  
Db 194 GLS-----YRVRVRAAOY-----NTWSEWSPSTKWNHSYREPEFHQLLLG 236  
Qy 212 VSEQHFGLSL-----YMNQVQGPKKR-----WHKNLTGQOI 244  
Db 237 VSVSIVILAVCLLCYVSTIKKEWDOI PNPASRLVAIIIDAOGSQWKEKRSRGOEP 296  
Qy 245 ITLNH-----TDLVPCLCIQVWPLEPDSVRTNICFPREDPR-----AHQNLWQAA 289  
Db 297 ACQPHWKNCLTKLLPCF-----LEHNMKRD-----EDPHKAAKEMPFGSGKSAWCPV 344  
Qy 290 RLRLTLQSW-----LLDAPCSLPAAEALCWAPGGDPCQPLVPLSWENVTVD 338  
Db 345 EISKTVL--WPESISVVRVCELFAPVECEEEV--EEEKSFC---ASPSSRDDFQE 397  
Qy 339 VNSSEKLQQLQCLWADSLGPKLDDVLLLETRG---PQDNRSICALPSPGCTS-----L 388  
Db 398 GREGIVARLTESLFDLLG-----EENGFCQDDMGESCLLPSPGSTSAHMPWDEF 448  
Qy 389 PSKASTRAARLGEYLLQDL-----OSGQCLQWLDDDL-----GALWACP 427  
Db 449 PSAGKPAAPPWGKEQPLHLEPSPASPPTQSPDNLTCTETPLVIAGNPAYRSFSNLSQSP 508  
Qy 428 MDKYL-----HKRWALVWLACILFAAALLKLDKDHAKGWLRLKQDV-RSGAAA 478  
Db 509 CPRELGPDPLLARHLEVEPEMPCV---PQISEPTTVPOPEPETWEQILRRNVLHQHAAA 565  
Qy 479 RGRAALLLSADDSGFERLVGALASALCQLPLRVAVDLWSRRLSAGQPVAFHQRROT 538  
Db 566 AP-----VSAPTSYQEFV-----HA----- 581  
Qy 539 LQEGG-----VVVLLFSPGAVA-----LCSEWLQDQVSGGAGHPHDAFRASLSCLVLD 587  
Db 582 VEQGTQASAVVGLGPPGGEAGYKAFSSLLASSAVSPEKCGFCASSGEEGYKP-----FQD 636  
Qy 588 FLQGRAPGSYGCADFRLHHPDAPALFRTVPVFT-----LPSQLPDFLG 632  
Db 637 LIPG-CPGD-----PAPVP-----VPLFTFGLDREPPRSPQSSHLFSSSPSEHLG 679  
Qy 633 A-----LQOPRAPRS-----GRQERAEQVSRALQPAL 660  
Db 680 LEPGEKVEDMPKRPPLPQSOATDPLVDSLGSIGVYSALTCHLCGHLKQCHQEDGGQTVM 739

QY 661 DS-----YFHPBCTPAPGRGVGPGAGP 682  
Db 740 ASPCCGCCCCDRSSPPTPLRAPDPSGGVP 770

## RESULT 13

A83583  
probable biotin synthesis protein BioC PA0503 [imported] - Pseudomonas aeruginosa (strain  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: A83583  
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337  
A:Accession: A83583  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <STO>  
A:Cross-references: GB:AE004487; GB:AE004091; NID:g9946361; PIDN:AAG03892.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0503

Query Match 2.9%; Score 106.5; DB 2; Length 274;  
Best Local Similarity 27.2%; Pred. No. 1.2; Indels 105; Gaps 22;  
Matches 92; Conservative 30; Mismatches 111

QY 371 PDNRSCLALEPSGCTSLPSK-----ASTRAAR-----LGEYLLQDLQSGCLO 414  
Db 2 PDDSSPL--LAPHGVAALPKQVAASFSAASYDAVLELQGVGESLLSALPEGSPR 59  
QY 415 LWDGDLGALWACP---MDKYIHRKALVWLACLLFAAALSILLKDKHAKGWLRLKQD 471  
Db 60 RW-VDLG---CGTGYFSRALERR-----FGAAEGLAV-----DIAEGMLR----- 95  
QY 472 VRSGAARGAALLLSADDSGFERLVGALASALCOLPLR-VAVDLMSRRLSQAQPVAV 530  
Db 96 ---HARAG-----GASHFEGDAE---RLPLRDGSCDL-----LFSLSAIFQW 132  
QY 531 -----FHAQRROTLOEGGVVLLFSPGAVALCSEWLDGVSFGCAHGPHDAFRASLSGV 584  
Db 133 CADLPAVLAFAARVLRPGG---VLAFLSLCVGTIGE-LRDSWRVVDGFGVHVRFRFA----- 184  
QY 585 LPDFLOGRAGSVYVACF---DRLLH-PD--AVPALFTVVFVFTLPOLDFLGLAQOQR 638  
Db 185 FADYLQHAAGSGLLPLTLRHEDRLHFPDLRLSLTHLKGALHNLNPRPDGL----- 237  
QY 639 APRSGRLOERAQVSRALQPALDSYFHPGCTPAPGRGV 676  
Db 238 ---TGR-----QRIRALVAAYERFRQEPGLPATYRVV 266

## RESULT 14

B75262  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: B75262  
R:White, O.; Eiken, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75262  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-679 <WHI>  
A:Cross-references: GB:AE002083; GB:AE000513; NID:g6460360; PIDN:AAF12085.1; PID:g646036

A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2546  
A:Map position: 1

Query Match 2.9%; Score 105.5; DB 2; Length 679;  
Best Local Similarity 26.1%; Pred. No. 4.7;  
Matches 81; Conservative 20; Mismatches 98; Indels 111; Gaps 17;

QY 438 LWWLACLLFA-----AALSILLKDKHAKGWLRLKQDVRSG-----AAARGAAL 484  
Db 71 LWWPLAALYGNRQLAGLLPLLPQQOFGVRRALLRADTTRRSFPDAPFPVAVRGRELA 130  
QY 485 LLYS-----ADDSGFERLVGALASALCOLPLRVAVDLWSRRLSAQ- 525  
Db 131 LAFAEHLAAGHHMGDVRPEHALVSDAGEVRLVGADYAL-----HLGRD 175  
QY 526 --GPV--AMFHAQRROTLOEG-----GVVLLFSPGAVALCSEWLDGVSFGP 569  
Db 176 FAGPVASAEYLPPEQRVGSATGAGADAESDAFLAVLFE-----LLGRHPYAGIOARG 231  
QY 570 -AHGPHDAFRASLSVLPDFLOGR--APGSYVGCACFDRLHDPVPAIF---RTVPVFTL 623  
Db 232 AAPGGEATAAGLFVDAPOAGFGRTAPGEWPFAL-----PPAVQALFVQAFVAVPVPR 286  
QY 624 PSOLPD-----FLGALQQ-----PRAPRSGRQLQERAEQVSRALQPALDSYFHPGCTPAPG--- 673  
Db 287 PS--PETWAALCALAAELVPCARRAG-----HWQVGLPCPCSCAA 325  
QY 674 --RGVGPAG 681  
Db 326 EREGAPSSG 335

## RESULT 15

B49555  
enhancer of split homolog R-esp2 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 26-May-2000  
C:Accession: B49555  
R:Schmidt, C.J.; Sladek, T.E.  
J. Biol. Chem. 268, 25681-25686, 1993  
A:Title: A rat homolog of the Drosophila enhancer of split (groucho) locus lacking  
A:Reference number: A49555; MUID:94064640  
A:Accession: B49555  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-741 <SCH>  
A:Cross-references: GB:L14463; NID:g294548; PIDN:AAC37640.1; PID:g294549  
A:Note: authors translated the codon TCC for residue 176 as Tyr, also Ala was found  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
F:497-530/Domain: WD repeat homology <WD1>  
F:583-616/Domain: WD repeat homology <WD2>  
F:665-698/Domain: WD repeat homology <WD3>  
F:706-739/Domain: WD repeat homology <WD4>

Query Match 2.9%; Score 105.5; DB 2; Length 741;  
Best Local Similarity 19.3%; Pred. No. 5.3;  
Matches 104; Conservative 54; Mismatches 147; Indels 235; Gaps 27;

QY 23 RLVGQDATHCSPGSLSCRLWDSILCLPGDIYPAG--PVLAP--THLOTETVLRCQKET 78  
Db 297 RTDAPTGTGNSNTPGLR-----PVPKPGVDPDLASSLRTPMAVPCPYPT 340  
QY 79 DCDLCLRVAVHLAVHGWEEPEDEKFGGAADSGVEPRNLAQVLSFQAYPTARC 138  
Db 341 P-----FGIVPHAGMNGELTSP-----CAAYAGL---HNISFQMSAAA----- 376  
QY 139 LLEVQVPAALVQFGSGSVVYDCFAALGSEVRINWSYTPRYEKELNHTQOLPALP---- 195  
Db 377 -----AAAAAAYGRS-----PVGFDPHHMRVPAIPPNL 407

```
QY 196 -----WLNYSADGDNTHVLVNSEEQHFGSLTYNOYQPPKPRMKNLTGP---- 242
Db 408 TGIPGKPAYSFHVVSADG-----OMQPYFFP--PDALIGPGIP 443
QY 243 -----QITLNTDLVPCL-----CLOVMPLEPDSVRTNICPFREDPRAHQ 283
Db 444 RHARQINTLNHGEVVCANTISNPTRHVYTGKGCYKVMWDITDPGNKSPVSQL--DCLNRD 501
QY 284 NLMQARL-----RLT-----TLQSWLIDAP-----CSLPAEALCMWRAPGDDPCQPLV 327
Db 502 NYIRSCRLPDGRTLIYVGEASTISIMDLAAPTPTKAKELTSSAPACY----- 549
QY 328 PPLSMENVTYVNSSEKIQLOBC-----LMADSLGPKLDVLLLETBPGPDNRSLCALE 381
Db 550 -----ALAI SPDSKVCFCSCSDGNIAVW-----DLHNTLYROFGHTDASCIDIS 596
QY 382 PSGCTSLPSKASTRARLGEYTLQDLOGGOCLOLMD--DDLGLMNCPM----- 428
Db 597 NDG-TKLMTGGLDNIVR-----SMDLREGROLQOHFTSQTSLGICPTGEMLAVGMENS 650
QY 429 -----DKY---IHKRNALWLAACILFAAALSLILLKKDHAKGWLRLKODVRSQ 475
Db 651 NVEVLHTKPKYQULHLE-----SCVLSLKFAGCKWF-----VRBG 688
```

Search completed: May 19, 2002, 14:11:38  
Job time: 6406 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 13:26:40 ; Search time 89.34 Seconds  
(without alignments)  
855.370 Million cell updates/sec

Title: US-09-608-918-10  
Perfect score: 3692  
Sequence: 1 MPVPWLLSLALGRSPVLS.....TPAPRGVGPAGPCAGDGT 688

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3692	100.0	688	22	Chimeric Zcytor14
2	3673.5	99.5	705	22	Human PRO polypept
3	3673.5	99.5	705	22	Human Interleukin
4	3673.5	99.5	705	22	Human PRO20040. H
5	3673.5	99.5	705	22	Chimeric Zcytor14
6	3606.5	97.7	675	22	Chimeric Zcytor14
7	3588	97.2	692	22	Human cytokine rec
8	3075	83.3	575	22	Human variant zcyt
9	1432	38.8	309	21	Human secreted pro
10	1305	35.3	267	22	Human membrane or
11	1300	35.2	332	22	Human protein sequ

12	1079	29.2	204	22	AAE10920	Human gene 12 enco
13	305.5	8.3	617	21	AAV76048	Murine skin cell p
14	305.5	8.3	617	22	AAB55987	Skin cell protein,
15	287.5	7.8	667	22	AAU04957	Human Interleukin
16	176	4.8	866	17	AAW04185	Human Interleukin
17	176	4.8	866	19	AAW61272	Human Interleukin
18	176	4.8	866	20	AAW92409	Human IL-17R prote
19	176	4.8	866	21	AAV99941	Human IL-17R prote
20	176	4.8	866	21	AAV97131	Human Interleukin
21	176	4.8	866	21	AAV97181	Human Interleukin
22	176	4.8	866	21	AAB03807	Human Interleukin
23	176	4.8	866	22	AAB62066	Human IL-17R (hCTL
24	176	4.8	866	22	AAV72754	Human Interleukin
25	166	4.5	864	17	AAW04184	Murine Interleukin
26	166	4.5	864	19	AAW61271	Mouse Interleukin
27	166	4.5	864	20	AAW92408	Murine IL-17R prot
28	166	4.5	864	21	AAV99935	Murine IL-17R prot
29	166	4.5	864	21	AAV97130	Murine Interleukin
30	166	4.5	864	21	AAV97180	Murine Interleukin
31	166	4.5	864	21	AAB03806	Murine Interleukin
32	166	4.5	864	22	AAB62060	Murine IL-17R poly
33	166	4.5	864	22	AAV72748	Murine Interleukin
34	147	4.0	330	21	AAV75947	Murine skin cell s
35	147	4.0	330	22	AAB55886	Skin cell protein,
36	126.5	3.4	1711	19	AAW70506	Osteostesticular pr
37	126.5	3.4	1711	19	AAW70507	Mutant osteostestic
38	116	3.1	5435	22	AAE10145	Streptomyces nous
39	115.5	3.1	2424	22	ABBS8934	Drosophila melanog
40	113	3.1	911	22	AAU69475	Human purified sec
41	113	3.1	4924	22	AAV70968	S. spinosa protein
42	113	3.1	4928	20	AAV93300	Spnd a polyketide
43	112.5	3.0	840	21	AAV93652	A mammalian solubl
44	112.5	3.0	840	21	AAV70028	Soluble Interleuki
45	111.5	3.0	586	22	AAU17979	Human immunoglobul

## ALIGNMENTS

### RESULT 1

ID	AAB61883	standard; Protein; 688 AA.
XX	AAB61883;	
DT	08-MAY-2001	(first entry)
XX	Chimeric Zcytor14 protein #1.	
DE	Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;	
KW	antiinflammatory; gene therapy; vaccine.	
XX	Homo sapiens.	
OS	WO200104304-A1.	
XX	18-JAN-2001.	
PD	30-JUN-2000; 2000WO-US18383.	
PF	07-JUL-1999; 99US-0348854.	
XX	(ZYMO ) ZYMOGENETICS INC.	
XX	Presnell SR, Burkhead SK, Pownder SL;	
PI	WPI; 2001-112618/12.	
XX	New polypeptide encoding a human cytokine receptor Zcytor14, for	
DR	treating inflammation e.g. rheumatoid arthritis -	
PT	Claim 2; Page 102-104; 112pp; English.	
XX		

CC The invention provides a new human cytokine receptor designated Zcytor14.  
 CC Zcytor14 can be expressed by standard recombinant methodology. The  
 CC encoding nucleic acid is useful for detecting the expression of a  
 CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
 CC used to screen biological samples in vitro for the presence of Zcytor14.  
 CC proteins, polypeptides and peptides having Zcytor14 activity can be  
 CC administered to a subject who lacks an adequate amount of this  
 CC polypeptide, for treating inflammation and conditions such as rheumatoid  
 CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14  
 CC antibodies) can be used to treat a subject who produces an excess of  
 CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
 CC Zcytor14 to a subject. The present sequence represents a chimeric  
 CC Zcytor14 protein.

XX Sequence 688 AA;

Query Match 100.0%; Score 3692; DB 22; Length 688;  
 Best Local Similarity 100.0%; Pred No. 0;  
 Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVPNFWLLSLGRSPVLSRLVGPQDATHCSPLGSLRLWSDILCLPGDIVPAPGPV 60  
 DB 1 mpvpnwllslgrspvlsrlvqpqathcsplgslcrldsdilclpgdivpapgpy 60

QY 61 LAPTHLOTELVRQKETCDICLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120  
 DB 61 lapthlqtelvrlcqketcdiclravhlavhghwepeedeekfggaadsgveeprnas 120

QY 121 LQAQVVLFOAYPTARCVLELVQVPAALVQFGOSGVSVVYDCFAALGSGEVRVWSYQPR 180  
 DB 121 lqaqvvlfoayptarcvlelvqvpaalvqfgosgvsvvydcfaalgsgsevrwswyqpr 180

QY 181 YKELNHTQOLPALPWLNVNSADGNDVHLVNVSEEHGFLSLXWNVQGGPKPRHKNLT 240  
 DB 181 yekelnhtqolpalpwlwnvnsadgndvhlvlnvseehgflslxwnvqggpkprwhknl 240

QY 241 GPQITLNTDLVPCLCIQWPLEPSVTRNICPFREDPRAHNLWQAARLLTLQSWL 300  
 DB 241 gpqitlntdlvpcliciqwplepsvtrnicpfredprahnlwqaarlrltlqswl 300

QY 301 LDAPCSLPAEALCWRAPGDCQPIVPLPSWENVTVDVNSSEKLOLOECILWADSLGPK 360  
 DB 301 ldapcslpaealcwragpdcqpilvplpswenvtvdvnssekloloeclwadslgpk 360

QY 361 DDVLLLETRGPQDNRSLCALFSGCTSLPSKASTRAARLGEYLLQDLQSGCQLQWDDDL 420  
 DB 361 ddvllletrgpqdnrslcalfsgctslpskastaarlgelyllqdlqsgcqlqwdddl 420

QY 421 GALWACPMCKYTHKRWALVWVLAACLLFAAALSLILLKDKHAKGWLRLKODVRSAAARG 480  
 DB 421 galwacpmdkythkrwalvwvllaacllfaaalslillkdkhagwlrllkdvrsaaarg 480

QY 481 RAALLYSADSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFHQRRTLQ 540  
 DB 481 raallysadsgferlvgalasalcqlplrvavdlwsrrelsaqgpvafhqrtrtlq 540

QY 541 EGGVVVLLFSPGAVALCSWLDGVSFGAHGPHDAFRASLSCVLPDFLOGRPGSVYGA 600  
 DB 541 eggvvllfspgavalcswldgvsfgahghpdafraslscvlpdflogrpgsvyga 600

QY 601 CFDRLLHPDAVPALFRTVPFTLPSPDLFLGALQOPRAPRSGRLQERAEQVSRALQPAL 660  
 DB 601 cfdrrllhpdavpalfrtvpftlpspdlflgalqopraprsgrlqeraeqvsrallpal 660

QY 661 DSYFHPGTPAPRGVGPAGPGAGDGT 688  
 DB 661 dsyfhpptpapgvgpapgagdggt 688

RESULT 2

AAU29322

ID AAU29322 standard; Protein; 705 AA.

XX AC AAU29322;  
 XX DT 18-DEC-2001 (first entry)  
 XX DE Human PRO polypeptide sequence #299.  
 XX KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200168848-A2.  
 XX PD 20-SEP-2001.  
 XX PF 28-FEB-2001; 2001WO-US06520.  
 XX PR 01-MAR-2000; 2000WO-US05601.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 03-MAR-2000; 2000US-187202P.  
 PR 06-MAR-2000; 2000US-186968P.  
 PR 14-MAR-2000; 2000US-189320P.  
 PR 14-MAR-2000; 2000US-189328P.  
 PR 15-MAR-2000; 2000WO-US06884.  
 PR 21-MAR-2000; 2000US-190828P.  
 PR 21-MAR-2000; 2000US-191007P.  
 PR 21-MAR-2000; 2000US-191048P.  
 PR 21-MAR-2000; 2000US-191314P.  
 PR 28-MAR-2000; 2000US-192655P.  
 PR 29-MAR-2000; 2000US-193032P.  
 PR 30-MAR-2000; 2000US-193053P.  
 PR 04-APR-2000; 2000US-194449P.  
 PR 04-APR-2000; 2000US-194647P.  
 PR 11-APR-2000; 2000US-195975P.  
 PR 11-APR-2000; 2000US-196000P.  
 PR 11-APR-2000; 2000US-196187P.  
 PR 11-APR-2000; 2000US-196690P.  
 PR 11-APR-2000; 2000US-196820P.  
 PR 18-APR-2000; 2000US-198121P.  
 PR 18-APR-2000; 2000US-198585P.  
 PR 25-APR-2000; 2000US-199397P.  
 PR 25-APR-2000; 2000US-199550P.  
 PR 25-APR-2000; 2000US-199654P.  
 PR 03-MAY-2000; 2000US-201516P.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 05-JUN-2000; 2000US-209832P.  
 PR 28-JUL-2000; 2000WO-US20710.  
 PR 22-AUG-2000; 2000US-0644848.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 08-NOV-2000; 2000WO-US30952.  
 PR 01-DEC-2000; 2000WO-US32678.  
 PR 20-DEC-2000; 2000WO-US34956.  
 XX (GETH ) GENENTECH INC.  
 PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
 XX WPI; 2001-602746/68.  
 DR N-PSDB; AAS46223.  
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
 PT presence of tumours, such as prostate and breast tumours, in mammals and  
 PT to screen for modulators of the compounds -  
 XX Claim 11; Fig 598; 774pp; English.  
 PS



Query Match	99.58%	Score 3673.5	DB 22	Length 705
Best Local Similarity	97.68%	Pred. No. 0		
Matches 688	Conservative 0	Mismatches 17	Gaps 1	

QY	1	MPVFFLLSLA	GRSPVVL	SERLTVG	PODATCS	GLSCRL	MDSDILC	PGGIVAPGV	60																																												
Db	1	mpvwfllslal	grspvvls	lerlvgr	pdaticsp	glscrl	mdsdilc	lpgdvlppgv	60																																												
QY	61	LAPHLQ	LELVLC	RCKETD	CDCL	RVAVHL	AHVHME	EBDEKEG	GAADSGVEE	RNAS	120																																										
Db	61	lapthlqel	vlrcqk	etcdcl	rlvavh	lawnh	weepede	ekfgaads	gvvee	rnas	120																																										
QY	121	LQAQV	VSFOAY	PTPAK	VILLE	QVPAAL	VQFQSG	SVYVDC	FEAALG	SEVATW	QPR	180																																									
Db	121	lqaqv	vsfay	ptarcv	llvq	vpaalv	qfsg	svsvydc	fleaalg	sevatw	qpr	180																																									
QY	181	YEKELN	HHQO	LPAL	PWLNV	SADGDN	VHVLV	LVN	SEOH	FECLSL	YMNVO	QPKPRM	HKNLT	240																																							
Db	181	yekeln	hqqp	lpa	lpw	lnvsad	gdnhv	lv	lnv	seehq	hglsl	ymvq	qppkpr	m	240																																						
QY	241	GPOIT	LHNHTD	LVPC	LCTQ	WMLPE	EDBSY	RTNICP	PERD	RAHQN	LMQAR	LILTL	QSWL	300																																							
Db	241	gpqitl	lnhtd	lvpc	lctq	wmlp	edbsv	rtnicp	perdra	hqn	lmaqar	lil	tlqswl	300																																							
QY	301	LDAPCS	LPAAEAL	CMRAP	GGDP	CPQLV	PPLS	MENTVD	-----VN	SSSE	343																																										
Db	301	ldapcs	lpaaeal	cwraps	gdcpql	lvprl	smentv	kdle	fp	llkghn	lcvq	vnse	360																																								
QY	344	KLQLO	ECIMAB	SLGRL	KNDV	LLER	RGQDN	SLCAL	EP	SGTSL	PSKRS	TAA	R	GETL	403																																						
Db	361	krlq	lqec	lwads	lgrl	kdv	lll	er	rgp	dln	slcal	eps	gts	lpsk	straet	lgeyl	420																																				
QY	404	LODOS	GGCLO	LMD	DD	LAL	MNAC	PM	DKYI	HK	KWAL	VW	LAC	L	FEAAL	SL	ILLK	KDH	ANG	463																																	
Db	421	lqdl	sggc	clq	wdd	l	al	wac	pndk	yl	hk	wal	v	la	cll	f	aaal	s	l	lll	k	dh	ang	480																													
QY	464	WLRL	KODV	RS	GAAR	GR	BA	LL	VS	AD	SG	FER	L	Y	GA	L	A	SAL	C	Q	LP	R	V	A	D	L	M	S	R	R	E	L	S	523																			
Db	481	wllrl	kqdv	rs	gaar	gr	ba	lla	ll	vs	ad	s	g	f	er	l	y	ga	l	a	s	a	l	c	q	p	r	v	a	d	l	m	s	r	r	e	l	s	540														
QY	524	AOGP	VA	W	HAOR	ROT	LO	B	G	V	V	LV	L	F	S	G	A	V	A	L	C	S	E	M	I	O	D	G	V	S	G	F	G	A	G	H	P	A	F	A	S	L	S	C	583								
Db	541	aogpva	w	ha	or	rot	l	b	g	v	v	lv	l	f	s	g	a	v	a	l	c	s	e	w	i	o	d	g	v	s	g	f	g	a	g	h	p	a	f	a	s	l	s	c	600								
QY	584	VL	PP	L	Q	R	A	G	R	A	G	S	V	Y	A	C	F	D	R	L	H	P	D	A	P	A	L	F	R	T	V	P	V	T	L	S	Q	L	P	D	F	L	G	A	Q	P	R	A	P	R	S	G	643
Db	601	vl	p	p	l	q	r	a	g	r	a	g	s	v	y	a	c	f	d	r	l	h	p	d	a	p	a	l	f	r	v	p	v	t	l	s	q	l	p	d	f	l	g	a	q	p	r	a	p	r	s	g	660
QY	644	RLQ	ER	A	E	O	V	S	R	A	L	O	P	A	L	D	S	F	H	P	E	T	P	A	P	A	G	S	V	G	A	P	G	A	D	G	T	668															
Db	661	rlqera	e	v	s	r	a	l	o	p	a	l	d	s	f	h	p	e	t	p	a	p	a	g	s	v	g	a	p	g	a	d	g	t	705																		

ID	AAU04956 standard; Protein: 705 AA.
AC	AAU04956;
AD	24-OCT-2001 (first entry)
AE	Human Interleukin 17 receptor, IL-17RH2.
AF	Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist;
AG	P8020040; DNA 164625-2890; systemic lupus erythematosus;
AH	Rheumatoid arthritis; osteoarthritis; diabetes mellitus;
AI	allergic disease; asthma; demyelinating disease;
AJ	degenerative cartilaginous disorder; transplantation associated disease
AK	Homo sapiens.
AL	Key
AM	Peptide
AN	Protein
AO	Region
AP	Modified-site
AQ	Region
AR	Modified-site
AS	Modified-site
AT	Modified-site
AV	Modified-site
AW	Modified-site
AX	Modified-site
AY	Region
AZ	Modified-site
BA	Modified-site
BB	Modified-site
BC	Modified-site
BD	Modified-site
BE	Modified-site
BF	Modified-site
BG	Modified-site
BH	Modified-site
BI	Modified-site
BJ	Modified-site
BK	Modified-site
BL	Modified-site
BM	Modified-site
BN	Modified-site
BO	Modified-site
BP	Modified-site
BQ	Modified-site
BR	Modified-site
BS	Modified-site
BT	Modified-site
BU	Modified-site
BV	Modified-site
BW	Modified-site
BX	Modified-site
BY	Modified-site
BZ	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site
CJ	Modified-site
CK	Modified-site
CL	Modified-site
CM	Modified-site
CN	Modified-site
CO	Modified-site
CP	Modified-site
CQ	Modified-site
CR	Modified-site
CS	Modified-site
CT	Modified-site
CU	Modified-site
CV	Modified-site
CW	Modified-site
CX	Modified-site
CY	Modified-site
CA	Modified-site
CB	Modified-site
CC	Modified-site
CD	Modified-site
CE	Modified-site
CF	Modified-site
CG	Modified-site
CH	Modified-site
CI	Modified-site

```

XX 23-DEC-1999; 99US-0172096.
PR 30-DEC-1999; 99WO-US31274.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 02-MAR-2000; 2000US-US05841.
PR 21-MAR-2000; 2000US-0191007.
PR 01-JUN-2000; 2000WO-US07532.
PR 02-JUN-2000; 2000WO-US15264.
PR 22-JUN-2000; 2000US-0213087.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 24-OCT-2000; 2000US-0242837.
PR 10-NOV-2000; 2000WO-US30873.
PR 28-NOV-2000; 2000US-0253646.
PR 01-DEC-2000; 2000WO-US32678.
XX
PA (GETH ) GENENTECH INC.
XX
PI Chen J, Filvaroff E, Pong S, Goddard A, Godowski PJ, Grimaldi CJ;
PI Gurney AL, Li H, Hillan KJ, Tumas D, Van Lookeren M, Vandlen RL;
PI Watanabe CK, Williams FW, Wood WI, Yansura DG;
XX
DR WPI; 2001-451708/48.
DR N-PSDB; AAS09515.
XX
PT Novel PRO polypeptides homologous to interleukin-17, useful for the
PT diagnosis and treatment of immune related disease e.g. rheumatoid
PT arthritis and diabetes .
XX
PS Claim 10; Fig 14; 18pp; English.
XX
CC The sequence is PRO20040 which is the human Interleukin 17 receptor,
CC IL-17RH2, encoded by DNA 164625-2890. A composition
CC containing ant/agonists to the PRO polypeptides or individual components
CC are useful for treating a mammal with an immune related disease, e.g.
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease, an autoimmune or immune-mediated skin
CC disease, contact dermatitis, an allergic disease e.g. food
CC hypersensitivity, asthma, a transplantation associated disease, or a
CC chronic inflammatory demyelinating polyneuropathy. Treating a
CC degenerative cartilaginous disorder comprises administering a PRO1031 or
CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous
CC examples of the diseases and disorders are given in the specification.
XX
SQ Sequence 705 AA;

```

```

Query Match 99.5%; Score 3673.5; DB 22; Length 705;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVLSRLVGPQDATHCSGLSRLWSDILCLPGDIVPAPGPV 60
DB 1 mpvpwflslalgrspvlsrlvqpqathcspslcrldwsdildclpgdivpappv 60

QY 61 LAPTHLQTELVRCKQKTCDCILCLRVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120
DB 61 lapthlqtelvrcqkctcdclclrvavhlavhghwepeedeekfggaadsgveeprnas 120

QY 121 LQAGVVLSPQAYPTARCVLLEQVPAALVQFGSGSVYDCFEALGSEVRWTSYQPR 180
DB 121 lqagvvlsfqayptarcvllvqpvaalvqfgsgsvvydcfealgservrwsytqpr 180

QY 181 YKELNHTQOLPALPWLNVASDGDVNLVNLVSEHQHFLSLYWNVOGPKPRWHKNLT 240
DB 181 ykelnhtqqlpalpwlvnvsadgdnvhlvlnvseehqfslslvynvqgppkprwhknlt 240

QY 241 GPQIITLHNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHQLNMQAARLRLITLQSWL 300

```

```

DB 241 gpqiitlhtdlvpcicqvwplesdvtrtnicpfredprahqnlwqaarlrlltqlswl 300
QY 301 LDAPCSLPAAALCWRAPOGDCQPLVPLSWENVTVD-----YNSSE 343
DB 301 ldapcslpaaaalcwraopgdpcqplvpplswenvvtvdvkvefpilkghpnlcvqvnsse 360
QY 344 KIQLOECLWADSLGPKLKDVLLETRGPDNRSLCALEPSGCTSLPSKASTAARLGEYL 403
DB 361 klqlqeciwadslgplkddvllletrgpdnrsilcalepsgctslpskastraarlgelyl 420
QY 404 LQDLSQGOCLOLWDDDLGALWACPMDKYTHKRWALVWLACLLFAAALSLLLLKKDHAKG 463
DB 421 lqdlsggqclqlwdddlgalwacpmdkyihkrwalvwlacllfaaalsllllkkdhakg 480
QY 464 WRLRLKQDVRSAAAARGAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELS 523
DB 481 wlrlkqdvrsaaaargaaalllysaddsgferlvgalasalcqlplrvavdlwsrrels 540
QY 524 AOGPVAVFHAQRQTLQEGVVVLLFSPGAVALCSEWLODGVSGAGHGPDAFASLSC 583
DB 541 aggpvavfhaqrqtlqegvvvllfsgpavalcsewldgvsgagahgphdafaaslsc 600
QY 584 VLPDFLOGRAPGSYVACFDRLLHPDAVPALFRTVPVFTLPSQLPDFLGAOQPRAPRS 643
DB 601 vlpdflograpsyvgacfdrlhpdavpalfrtvpvftcpsqlpdfigalqgpraprs 660
QY 644 RLQRAEQVSRALQALPDALSYFHPGTPAPGRGVGPGACGAGDGT 688
DB 661 rlqraeqvstralqaldsyfhpptpapgryvgpgagpgagdggt 705

```

```

RESULT 4
AAB87606
ID AAB87606 standard; Protein; 705 AA.
XX
AC AAB87606;
XX
DT 15-MAY-2001 (first entry)
XX
DE Human PRO20040.
XX
KW Human; PRO protein; mapping.
XX
OS Homo sapiens.
XX
PN WO200116318-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US23328.
XX
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 07-DEC-1999; 99US-0169495.
PR 09-DEC-1999; 99US-0170262.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-019397.
PR 22-MAY-2000; 2000WO-US14042.
PR 05-JUN-2000; 2000US-0209832.
XX
PA (GETH ) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR N-PSDB; AAF92138.

```

XX Eighty four nucleic acids encoding PRO polypeptides, useful in  
 PT molecular biology, including use as hybridization probes, and in  
 PT chromosome and gene mapping.  
 XX  
 PS  
 XX Claim 12; Fig 162; 278bp; English.

CC The present sequence is a human PRO polypeptide (secreted and  
 CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or  
 CC anti-PRO antibodies are useful for preparation of a medicament useful in  
 CC the treatment of a condition which is responsive to the PRO protein,  
 CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be  
 CC employed as molecular weight markers for protein electrophoresis. The PRO  
 CC coding sequence has applications in molecular biology, including use as  
 CC hybridisation probes, and in chromosome and gene mapping.  
 CC  
 XX  
 XX

Sequence 705 AA:

Query Match 99.5%; Score 3673.5; DB 22; Length 705;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVLSLERLVGPQDATHGSPGSLCRLMSDITLCPGDIVAPGPV 60  
 DB 1 mpvpwflslalgrspvlslerlvpgqdaathgspgslcrlmsdiltcpgdivapgpv 60  
 QY 61 LAPTHLQTELVLRCOKETDCDLCRLVAHVLAHGHWEPEDEKFGGAADGVEEPRNAS 120  
 DB 61 lapthqltelvlrcoketdcdlclrvahlvahghwepeedeekfggsadsgveepnas 120  
 QY 121 LQAQVVLSPQAYPTARCVLLEVOVPALVQFGSGSVVYDCFEALGSEVRIMSTQPR 180  
 DB 121 lqaqvvlspqayptarcvlllevqpaaivqfgsgsvvydcfealgservimstqpr 180  
 QY 181 YEKELNHTQOLPALPMLNVSADGDNVHLVNVSEOHFGSLVYNOVGPPKPRMKNLT 240  
 DB 181 yekelnhtqolpalpmlnvsadgdnvhlvnvseohfgslvynovgppkprwmknlc 240  
 QY 241 GPQITLNLHNDLVPCICIQWPLEPDSVRNTPCFREDPRAHONLMOAARLRLTLQSWL 300  
 DB 241 gpqitlntldlvpciciqwplepdsvrtnicpfredprahnlwgaarlrlltlqswl 300  
 QY 301 LQAPCSLPAPALCMWRAGGDCPOPLVPPLSMENVTVD-----VNSSE 343  
 DB 301 ldapcslpaeaalcmwragsgdcpcplvpplsmenvtvdvlefpllkxhpnlcvgvnase 360  
 QY 344 KIQLOECIMADSLGRLKDVLLLETRGPDNRSLCALEPSGCTSLPSKASRAARLIGYL 403  
 DB 361 kIQlgeclwadsIgpLkddvlllctrgpdnrsIcaIePsgctslpskastraaRlIgeYl 420  
 QY 404 LQDLQSGOCLQWLMDDDLGLALMACPMKDYIHKRMALVWLACLIFAAALSLILLKDKHAKG 463  
 DB 421 lqdlqsgqclqldwddlglalwacpmkdyIhkrmalvwlacllfaaalsllllkkdhakg 480  
 QY 464 WELRLKQDVNSGAARRAALLLYSADSGFERLVGALASLQCLPLFVAVDMSRREL 523  
 DB 481 wllllkqdvnsGaaraaRlaalllySaddsgferlvgaIasalcqplrlvavdlwrrrels 540  
 QY 524 AGCPVAMFHHOROTLOEGGVVLLFSGVALCSEMIQDQVSGGAGHPDAFASISC 583  
 DB 541 aggpvaafhaqrrqcllqeggvvlllfsGvalcsewIqdgvsGpagaRphdaIfaasIsc 600  
 QY 584 VLDPFLQGRAGSVYACFRLRLHPDAVPALFRTVPVFTLSQLPDEFGLAQOFRAPPSG 643  
 DB 601 vlpdfllqgragsvyagcfdrllhpdevpalfrtvpvftlpsqldpdefglagqprapsG 660  
 QY 644 RLQERAEQVSRAALQPALDSYFHPGTPAPRGVSGVAGAGAGDGT 668  
 DB 661 rlqereaeqvsraIqpaldsyfhpgtpaprgvsgvsgagagdggt 705

RESULT 5

AAB61884  
 ID AAB61884 standard; Protein; 705 AA.  
 XX  
 AC AAB61884;  
 XX  
 XX

DT 08-MAY-2001 (first entry)

DE Chimeric Zcytor14 protein #2.

KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
 KW antiinflammatory; gene therapy; vaccine.

OS Homo sapiens.

PN WO200104304-A1.

PD 18-JAN-2001.

PF 30-JUN-2000; 2000WO-US18383.

PR 07-JUL-1999; 99US-0348854.

PA (ZYMO ) ZYMOGENETICS INC.

PI Presnell SR, Burkhead SK, Powderer SL;

DR WPI; 2001-112618/12.

PT New polypeptide encoding a human cytokine receptor zcytor14, for  
 treating inflammation e.g. rheumatoid arthritis -

PS Claim 2; Page 105-107; 112pp; English.

CC The invention provides a new human cytokine receptor designated zcytor14.  
 CC zcytor14 can be expressed by standard recombinant methodology. The  
 CC encoding nucleic acid is useful for detecting the expression of a  
 CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be  
 CC used to screen biological samples in vitro for the presence of zcytor14.  
 CC Proteins, polypeptides and peptides having zcytor14 activity can be  
 CC administered to a subject who lacks an adequate amount of this  
 CC polypeptide, for treating inflammation and conditions such as rheumatoid  
 CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14  
 CC antibodies) can be used to treat a subject who produces an excess of  
 CC zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
 CC zcytor14 to a subject. The present sequence represents a chimeric  
 CC zcytor14 protein.  
 CC  
 XX  
 XX

Sequence 705 AA:

Query Match 99.5%; Score 3673.5; DB 22; Length 705;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 688; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVLSLERLVGPQDATHGSPGSLCRLMSDITLCPGDIVAPGPV 60  
 DB 1 mpvpwflslalgrspvlslerlvpgqdaathgspgslcrlmsdiltcpgdivapgpv 60  
 QY 61 LAPTHLQTELVLRCOKETDCDLCRLVAHVLAHGHWEPEDEKFGGAADGVEEPRNAS 120  
 DB 61 lapthqltelvlrcoketdcdlclrvahlvahghwepeedeekfggaadsgveepnas 120  
 QY 121 LQAQVVLSPQAYPTARCVLLEVOVPALVQFGSGSVVYDCFEALGSEVRIMSTQPR 180  
 DB 121 lqaqvvlspqayptarcvlllevqpaaivqfgsgsvvydcfealgservimstqpr 180  
 QY 181 YEKELNHTQOLPALPMLNVSADGDNVHLVNVSEOHFGSLVYNOVGPPKPRMKNLT 240  
 DB 181 yekelnhtqolpalpmlnvsadgdnvhlvnvseohfgslvynovgppkprwmknlc 240  
 QY 241 GPQITLNLHNDLVPCICIQWPLEPDSVRNTPCFREDPRAHONLMOAARLRLTLQSWL 300  
 DB 241 gpqitlntldlvpciciqwplepdsvrtnicpfredprahnlwgaarlrlltlqswl 300

QY 301 LDAPCSLPARAALCWRAPGDPQCPPLVPPLSWENVTVD-----VNSSE 343  
|||||  
Db 301 ldapcslpaeaalcwrappgdpqcpplvpplswenvtdkvlefpllkghpnlcvqvnss 360  
QY 344 KIQLOECLWADSLGPKDDVILLLETRGPQDNRSICALPEPGCTSLPSKASTRAARLGEYL 403  
|||||  
Db 361 kqlqecclwadsigplkddvillletrgpqdnrsicalpepgctslpskastaaralgeyl 420  
QY 404 LQDLSQOCLQWLDDDLGALWACPMCKY IHKRWALVWLACLLFAAALSILLKKDHAKG 463  
|||||  
Db 421 lqdlsgqclqldddlgalwacpmckyihkrwalvwlacllfaaalsilllkkdhakg 480  
QY 464 WRLRLKQDVRSGAARGAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRRLS 523  
|||||  
Db 481 wrllkqdvrsagaargaalysadsgferlvgalasalcolplrvavdlwsrrrls 540  
QY 524 AQGPVAFHQAORQTLOEGVVVLLFSPGAVLCEWLQDGVSGPGAHGPHDAFRASLSC 583  
|||||  
Db 541 aggpvafhaqrqtllqeggvvllfsgpavalcsewlqdgvsppgahghphdafrafrasc 600  
QY 584 VLPDFLQGRAPGSVVGACFDRLHDPDAVPALFRVTPVFTLPSQLPDFLQALQOPRPRSG 643  
|||||  
Db 601 vipdfllqgrapgsyvgacfdrlhdpdvpalftrtpvftlpsqlpdfllqalqqprprsg 660  
QY 644 RLQERAEQVSRALQPALDSYFHPGTPAPGRGVGPGAGGAGDGT 688  
|||||  
Db 661 rlqeraeqvsralqpaldsyfhpptpapgrgvgpgagpgagdgdt 705

RESULT 6  
AAB61885  
ID AAB61885 standard; Protein; 675 AA.  
XX AAB61885;  
AC AAB61885;  
XX  
DT 08-MAY-2001 (first entry)  
DE Chimeric Zcytor14 protein #3.  
XX  
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
KW antinflammatory; gene therapy; vaccine.  
OS Homo sapiens.  
XX  
PN WO200104304-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 30-JUN-2000; 2000WO-US18383.  
XX  
PR 07-JUL-1999; 99US-0348854.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
PI Presnell SR, Burkhead SK, Powder SL;  
PI  
XX WPI; 2001-112618/12.  
XX  
PT New polypeptide encoding a human cytokine receptor Zcytor14, for  
PT treating inflammation e.g. rheumatoid arthritis -  
PS  
PS Claim 2; Page 107-109; 112pp; English.  
XX  
CC The invention provides a new human cytokine receptor designated Zcytor14.  
CC Zcytor14 can be expressed by standard recombinant methodology. The  
CC encoding nucleic acid is useful for detecting the expression of a  
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
CC used to screen biological samples in vitro for the presence of Zcytor14.  
CC Proteins, polypeptides and peptides having Zcytor14 activity can be  
CC administered to a subject who lacks an adequate amount of this  
CC polypeptide, for treating inflammation and conditions such as rheumatoid  
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14

CC antibodies) can be used to treat a subject who produces an excess of  
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
CC Zcytor14 to a subject. The present sequence represents a chimeric  
CC Zcytor14 protein.

XX Sequence 675 AA;

Query Match 97.7%; Score 3606.5; DB 22; Length 675;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 675; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPVPWFLLSLALGRSPWLSLERLVGPQDATHCSPLGSLWSDILCLPGDIVPAGPV 60  
|||||  
Db 1 mpvpwfllslalgrspwlslerlvgpqdathtcsplgslwdsilclpgdivpapgvp 60  
QY 61 LAPTHLQTELVLRQKETDCDCLRVAVHLAVHGHWEPEDEKFGAADSVEEPRNAS 120  
|||||  
Db 61 lapthlqtevlvrqketdcdclrvavhlavhghweepedeekfgaadsveeprnas 120  
QY 121 LQAQVVLSPQAYPTARCVLLEVOVPAALVQFGOSVGVVYDCFEALGSEVRITWSTQPR 180  
|||||  
Db 121 lqaqvvlspqayptarcvlllevopaalvqfgosvsvvycfealgservitwsyqpr 180  
QY 181 YKELNHTQQLPALPWLNVSDGDNVHLVNLVSEEHFGLSLYMNVOGPKPRWHKNLT 240  
|||||  
Db 181 ykelnhtqqlpalpwlrvnsadgnvhlvlnvseeqhfglslvnmvgpkrwhknlt 240  
QY 241 GPQITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLTQSWL 300  
|||||  
Db 241 gpqitlntdlvpcicigwplepdsvrtnicpfredprahonlwqaarlrltlqswl 300  
QY 301 LDAPCSLPAAALCWRAPGDPQCPPLVPPLSWENVTVDVNSSEKIQLOECLWADSLGPK 360  
|||||  
Db 301 ldapcslpaeaalcwrappgdpqcpplvpplswenvtdvnsseklqleclwadsigplk 360  
QY 361 DVVLLLETRGPQDNRSICALPEPGCTSLPSKASTRAARLGEYLILQDLQSQCLQWLDDDL 420  
|||||  
Db 361 dvvllletrgpqdnrsicalpepgctslpskastaaralgeylilqdlsgqclqldddd 420  
QY 421 GALWACPMCKY IHKRWALVWLACLLFAAALSILLKKDHAKWLRLLKODVRSAGAAARG 480  
|||||  
Db 421 galwacpmckyihkrwalvwlacllfaaalsilllkkdhak-----aaarg 467  
QY 481 RAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRRLSACQPVAFHQAORRQTLO 540  
|||||  
Db 468 raallysadsgferlvgalasalcolplrvavdlwsrrrelsaqpvafhaqrqtllq 527  
QY 541 EGGVVVLLFSPGAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGA 600  
|||||  
Db 528 eggvvvllfsgpavalcsewlqdgvsppgahghphdafrafrascvlpdfllqgrapgsyvga 587  
QY 601 CFDRLLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRPRSGRLQERAEQVSRALQPAL 660  
|||||  
Db 588 cfdrllhpdavpalftrtpvftlpsqlpdfllqalqoprprsgrlqeraeqvsralqpal 647  
QY 661 DSYFHPGTPAPGRGVGPGAGGAGDGT 688  
|||||  
Db 648 dsyfhpptpapgrgvgpgagpgagdgdt 675

RESULT 7

AAB61880  
ID AAB61880 standard; Protein; 692 AA.  
XX  
AC AAB61880;  
XX  
DT 08-MAY-2001 (first entry)  
XX  
DE Human cytokine receptor Zcytor14.  
XX  
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
KW antinflammatory; gene therapy; vaccine.

XX OS Homo sapiens.  
XX PN WO200104304-A1.  
XX PD 18-JAN-2001.  
XX PF 30-JUN-2000; 2000WO-US18383.  
XX PR 07-JUL-1999; 99US-0348854.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Presnell SR, Burkhead SK, Pownder SL;  
XX DR WPI; 2001-112618/12.  
XX DR N-PSDB; AAC85027.  
XX PT New polypeptide encoding a human cytokine receptor Zcytor14, for  
XX PT treating inflammation e.g. rheumatoid arthritis -  
XX PS Claim 2; Page 2; 112pp; English.  
XX CC The invention provides a new human cytokine receptor designated Zcytor14.  
XX CC Zcytor14 can be expressed by standard recombinant methodology. The  
XX CC encoding nucleic acid is useful for detecting the expression of a  
XX CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
XX CC used to screen biological samples in vitro for the presence of Zcytor14.  
XX CC proteins, polypeptides and peptides having Zcytor14 activity can be  
XX CC administered to a subject who lacks an adequate amount of this  
XX CC polypeptide, for treating inflammation and conditions such as rheumatoid  
XX CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14  
XX CC antibodies) can be used to treat a subject who produces an excess of  
XX CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
XX CC Zcytor14 to a subject. The present sequence represents the human  
XX CC cytokine receptor Zcytor14.  
XX SQ Sequence 692 AA;

Query Match 97.2%; Score 3588; DB 22; Length 692;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 675; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

Qy 1 MPVPWFLSLALGRSPVLSRLVLPQDATHCSPLGSLRSLWSDILCLPGDIVPAPGPV 60  
Db 1 mpvpwflslalgrspvlsrlvlpqdgathcspslgsrlwdsdildclpgdivpapg 60  
Qy 61 LAPTHLOTELRLRCOKETDCDCLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPRNAS 120  
Db 61 laphlqltelvircqketdcdclrlvavhlavhghweepekfggaadsgveeprnas 120  
Qy 121 LQAOVVLSFOAYPTARCVLLEVQVPAALVQFGSGSVVYDCFEAALGSEVRWISYTOPR 180  
Db 121 lqaqvvlslfayptarcvllvqvpaalvqfgsgsvvydcfeaalgservrwlswytcpr 180  
Qy 181 YEKELNHTQQLPALPWLNVYADGDNVHLVNVSEEHFGLSLYWNQVQPKPRWHKNLT 240  
Db 181 yekelnhtqqlpalpwnvsadgdnvhlvlnvseeqhfglslwnvqvgpkprwhknlt 240  
Qy 241 GPQIITLNTHTDVLPCIQVPLEPDSVRTNICPFREDPRAHNLQMAARLLTLOSWL 300  
Db 241 gpiitlnthtdvlpciqvplepdsvrtnicpfredprahnlwqaarlrlltlqswl 300  
Qy 301 LDAPCSLPAAALCWAPGDDPCQPLVPPLSWENVTVD-----VNSSE 343  
Db 301 ldapcslpaaalcwapgdgpcqplvpplswenvtdkvleflklxghnplcvqvnasse 360  
Qy 344 KLQLECLWADSLGPLKDDVLLLETRGPQDNRSICALPEPGCTSLPSKASTRAARLGEYL 403  
Db 361 klqlqecslwadslgplkddvllletrgpqdnrsicalpepgctslpskastaarlgeyl 420  
Qy 404 LQDLQSGQCQLQWDDDLGALWACPMDDKYIHKRWALVWLACLLFAAALLSLILLKKDHAKG 463

Db 421 lqdlsggcqlqwdldlgalwacpmdkyihkrwalwlacllfaaalsiillllkkdhak- 479  
Qy 464 WLRLKQDVRSGAAAGRAALLLYSADDSGFERLVGALASALCQPLRVAVDLSRRRELS 523  
Db 480 -----aaagraalllysaddsgferlvgalasalcqiplrvavdlwrrrels 527  
Qy 524 AOGPVAFHAQRROTLOEGGVVVLLFSPCAVALCSEWLDGVSGPCAHGPHDAFRASLSC 583  
Db 528 aqgpvawfhaqrrotlqeggvvvllfsgpavalcsewldqvgsgpahghpdafraslsc 587  
Qy 584 VLPDFLOGRAPGSYVGACFDRLHDPDAVPALFRTVPVFTLPOLPDFLGCALQOAPRPSG 643  
Db 588 vlpdflograpgsyvgacfdrlhdpavpalftvpvftlpqldfqlgqlqaprasg 647  
Qy 644 RLQERAEQVSRALQPALDSYFHPPTGTPAPGRGVGPGAGGAGDGT 688  
Db 648 rlqeraeqvsralqpaldsyfhpptgtpapgrgvpgagpgagcgt 692

RESULT 8  
AAB61881  
ID AAB61881 standard; Protein; 575 AA.  
AC AAB61881;  
XX 08-MAY-2001 (first entry)  
DE Human variant Zcytor14 protein Zcytor14-1.  
XX Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
XX antiinflammatory; gene therapy; vaccine; variant; Zcytor14-1.  
XX Homo sapiens.  
XX WO200104304-A1.  
XX 18-JAN-2001.  
XX 30-JUN-2000; 2000WO-US18383.  
XX 07-JUL-1999; 99US-0348854.  
XX (ZYMO ) ZYMOGENETICS INC.  
XX Presnell SR, Burkhead SK, Pownder SL;  
XX WPI; 2001-112618/12.  
XX N-PSDB; AAC85029.  
XX New polypeptide encoding a human cytokine receptor Zcytor14, for  
XX treating inflammation e.g. rheumatoid arthritis -  
XX Disclosure; Page 2-3; 112pp; English.  
XX The invention provides a new human cytokine receptor designated Zcytor14.  
XX Zcytor14 can be expressed by standard recombinant methodology. The  
XX encoding nucleic acid is useful for detecting the expression of a  
XX Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
XX used to screen biological samples in vitro for the presence of Zcytor14.  
XX Proteins, polypeptides and peptides having Zcytor14 activity can be  
XX administered to a subject who lacks an adequate amount of this  
XX polypeptide, for treating inflammation and conditions such as rheumatoid  
XX arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14  
XX antibodies) can be used to treat a subject who produces an excess of  
XX Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
XX Zcytor14 to a subject. The present sequence represents a variant of  
XX the human cytokine receptor Zcytor14, designated Zcytor14-1. This  
XX variant is a truncated form of the receptor polypeptide and lacks  
XX amino acid residues 1-113 of Zcytor14.  
XX Sequence 575 AA;

Query Match	83.3%;	Score 3075;	DB 22;	Length 575;
Best Local Similarity	100.0%;	Pred. No. 1.1e-274;		
Matches 575;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 114	EEPRNASLQAQVVLSTQAYTARCVLLEVOVPAALVFGVSGVVDVDFCAAALGSEVRI	173		
Db 1	eeprnaslqaqvvlstqaytarcvllevdpaalvfgvsgvvdvdfcaaalgsevril	60		
QY 174	WSVTPRYEKELNHTQOLPALPLNVSADGDNVHLVNLVSEEQHFGLSLYWNOVQGPKP	233		
Db 61	wsytpryekelnhtqqlpalplwnvsadgdnvhlvlnvseeqhfghslywnvqggppkp	120		
QY 234	RHKNLTGPQITLNTHTDVLVPCICIQWPLEPDSVRTNICPFREDPRHAQNLAARLRL	293		
Db 121	rwkhnltgpqitlnhtdlvpcicqvwplepdsrvtrnicpfredprahnqlwqaarlrl	180		
QY 294	LTLQSWLLDAPCSLPAEALCWAPGDPQCPVPPPLSWENVTVDVNSSEKQLQOECLWA	353		
Db 181	ltlqswlldapcslpaealcwragdpqcpvppplswenvtvdsseklqlqecclwa	240		
QY 354	DSLGLPKDVLLETGPDNRSLCALEPSGCTSLPSKASTRAARLGEYLLQDLSGOCL	413		
Db 241	dslgplkddvllletrgpqnrslicalepgsctslpskastraarlgeyllqdlqsgqcl	300		
QY 414	QLWDDDLGALWACPMQKYIHKRWALVWLACLLFAAALSLLILLKKHAKGWLRLKQDVR	473		
Db 301	qlwdddlgalwacpmdkyihkrwalvwlacllfaaalsllllkkhakgwrlrkqdv	360		
QY 474	SGAAARGRAALLYSADDSGFERLWALASALCOLPURVAVDLMSRRELSAQGPVAFHA	533		
Db 361	sgeaargraallysaddsgferlvgalasalcolpurvavdlmsrrelsaqgpvafha	420		
QY 534	QRROTLEGGVVVLLPSGAVALCSEWLQDVSQSGAGHPHDAFRASLSCVLPDFOGRA	593		
Db 421	qrrotleggvvvllfsgavalcsewldqvsqsgaghpghdafraslscvlpdflgra	480		
QY 594	PGSYGACFDRLHPDAVPALFRTVPVFTLPSQLPDFLQALQOPRAPRSGRQLRAEQVS	653		
Db 481	pgsygacfdrlhpdavpalfrtvpvftlpsqlpdflqalqqpraprsgrlqeraeqvs	540		
QY 654	RALQPALDSYFHPGTPAPGRGVPAGGAGDGT 688			
Db 541	ralqpaldsyfhpptpapgvgpaggagdgdt 575			
RESULT 9				
AA76143				
ID	AA76143 standard; Protein: 309 AA.			
AC	AA76143;			
DT	23-MAR-2000 (first entry)			
XX	Human secreted protein encoded by gene 20.			
DE				
XX	Human; secreted protein; cancer; tumour; developmental abnormality;			
KW	foetal deficiency; blood disorder; immune system disorder; inflammation;			
KW	autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;			
KW	schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;			
KW	atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;			
KW	digestive disorder; endocrine disorder; infection; AIDS; leukaemia;			
KW	therapy; chromosome 3.			
XX				
OS	Homo sapiens.			
XX				
PN	WO9958660-A1.			
XX				
PD	18-NOV-1999.			
XX				
PF	06-MAY-1999; 99WO-US09847.			
XX				
PR	12-MAY-1998; 98US-0085093.			

PR 12-MAY-1998;	98US-0085094.
PR 12-MAY-1998;	98US-0085105.
PR 12-MAY-1998;	98US-0085180.
PR 18-MAY-1998;	98US-0085906.
PR 18-MAY-1998;	98US-0085920.
PR 18-MAY-1998;	98US-0085921.
PR 18-MAY-1998;	98US-0085922.
PR 18-MAY-1998;	98US-0085923.
PR 18-MAY-1998;	98US-0085924.
PR 18-MAY-1998;	98US-0085928.
PR 18-MAY-1998;	98US-0085925.
PR 18-MAY-1998;	98US-0085927.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI	Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI	Lafleur DW, Endress GA, Ebner R;
XX	WPI; 2000-062296/05.
DR	N-PSDB; AAZ65269.
XX	New isolated human genes and the secreted polypeptides they encode,
PT	useful for diagnosis and treatment of e.g. cancers, neurological
PT	disorders, immune diseases, inflammation or blood disorders
XX	Claim 11; Page 372-373; 475pp; English.
XX	AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC	AAZ76124 to AAZ76223 are the secreted proteins encoded by the 97 human
CC	genes. The gene encoding this protein was found to be on chromosome 3.
CC	The genes and their corresponding secreted polypeptides are
CC	useful for preventing, treating or ameliorating medical conditions,
CC	e.g. by protein or gene therapy. Also pathological conditions can be
CC	diagnosed by determining the amount of the new polypeptides in a sample
CC	or by determining the presence of mutations in the new genes. Specific
CC	uses are described for each of the 97 genes, based on which tissues they
CC	diagnose or treatment of cancer, tumours, developmental abnormalities
CC	and foetal deficiencies, blood disorders, diseases of the immune system,
CC	autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC	disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC	disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC	disorders, digestive/endocrine disorders, infections and AIDS. The
CC	polypeptides are also useful for identifying their binding partners.
CC	The sequences shown in AAY76224 to AAY76424 represent fragments of the
CC	secreted proteins.
XX	Sequence 309 AA;
SQ	

Query Match	38.8%;	Score 1432;	DB 21;	Length 309;
Best Local Similarity	90.9%;	Pred. No. 1.8e-123;		
Matches 270;	Conservative 0;	Mismatches 11;	Indels 16;	Gaps 2;
QY 1	MPVPWFLLSALGRSPVLSRLVGPQDATHCSGLSCLRLWDSILCLPGDIVPAGPV 60			
Db 1	mpvpwflslalgrspvlsrlrvpqpqathcsprlscrlwdsilclpgdivpaggv 60			
QY 61	LAPTHLQTELVLRCOKETDCDCLRLRVAVHLVHGHWEPEDEKFGGAADSGVEEPRNAS 120			
Db 61	laphlqtelvlrcqketdcdclrvxvhlavhghwepeedeekfggaadlsgveeprnas 120			
QY 121	LQAQVLSFQAYPTARCVLLEVOVPAALVFGVSGVVDVDFCAAALGSEVRISWSTOPR 180			
Db 121	lqaqvlsfayptarcvllvqvpaalvfgvsgvsvvydcfeaalgsvevrilswstqpr 180			
QY 181	YEKELNHTQOLP-----ALPWLNVSDAGDNVHLVNLVSEEQHFGLSLYWN 225			
Db 181	yekelnhtqqlpdcrglevnswpsicwlpwnlnvsadgdnvhlvlnvseeqhfghslywn 240			
QY 226	QVQGGPKPRWHKNTLGPQITLNTHTDVLVPCICIQWPLEPDSVRTNICPFREDPRAH 282			
XX				



CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antinaemic; antiagregant; haemostatic; vulnerary;  
 CC anticancer; osteopathic; dermatological; antiallergic; antiasthmatic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.  
 XX  
 SQ Sequence 332 AA;

Query Match 35.2%; Score 1300; DB 22; Length 332;  
 Best Local Similarity 86.4%; Pred. No. 3e-111;  
 Matches 248; Conservative 5; Mismatches 12; Indels 22; Gaps 3;  
 QY 1 MPVPFLLSLALGRSPVLSRLERLVGPQDATHGSPGLSRLWDSIDLCLPGDIVPAGPV 60  
 DB 2 MPVPFLLSLALGRSPVLSRLERLVGPQDATHGSPGLSRLWDSIDLCLPGDIVPAGPV 61  
 QY 61 LAPTHLQTELVLRCQETDCDLCLRVAVHLAVHGWEEPEDEKFGGAAGSGVEEPRNAS 120  
 DB 62 lapthlqtelvlrcqetdcdclrvavhlavhgw eepeedeekfggaadsgveeprnas 121  
 QY 121 LQAQVVLFSQAYTARCVLLEVOVPAALVQFGSGVSVVDCFAALGSEVRISYQTPR 180  
 DB 122 lqaqvvlfsqaytarcvlllevvpaalvqfgsgsvvvdcfaaalgsevr isyqtpr 181  
 QY 181 YKELNHTQQLP-----ALPWLNVSDGDNVHLVNVSEOHFGLSLYWN 225  
 DB 182 ykelnhtqqlpcrcglevwnspscwlpwlnvsadgdnvhlvlnvseeqhfslslywn 241  
 QY 226 QVGGPPKPRWKNLTGP---QITLNTHTDLVPCLCIQVWPLEPDSVR 269  
 DB 242 qvggppkprwknlvpppsqv----hshcrpcclckadvpyqrgslk 284

RESULT 12  
 AAE10920  
 ID AAE10920 standard; Protein; 204 AA.  
 XX  
 AC AAE10920;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human gene 12 encoded immune system-related protein HADEX14.  
 XX  
 KW Human: immune system-related protein; allergy; rheumatoid arthritis;  
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;  
 KW diabetes mellitus; arrhythmia; wound healing; ischaemic lesion; AIDS;  
 KW Acquired Immune Deficiency Syndrome; viricide; hepatotropic; vasotropic;  
 KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;  
 KW wound healing; cell proliferation; skin aging; endocrine disorder;  
 KW food preservative.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 50..55  
 FT /label= Immunogenic\_epitope  
 FT 151..168  
 FT /label= Immunogenic\_epitope  
 FT  
 XX

PN WO2000166722-A1.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 07-MAR-2001; 2001WO-US07260.  
 XX  
 PR 08-MAR-2000; 2000US-187873P.  
 PR 11-AUG-2000; 2000US-224367P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;  
 PI Gruber JR, Endress GA, Ruben SM;  
 XX  
 DR WPI; 2001-589939/66.  
 DR N-PSDB; AAD18278.  
 XX  
 PT Novel isolated immune system-related polypeptide useful for treating  
 PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,  
 PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and  
 PT viral hepatitis  
 XX  
 PS Claim 11; Page 312-313; 315pp; English.  
 XX  
 CC The invention relates to human immune system-related protein and their  
 CC DNA. Human immune-system related protein and DNA are useful for  
 CC preventing, treating or ameliorating a medical condition in a mammalian  
 CC subject, for diagnosing, preventing or treating immune system-associated  
 CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory  
 CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders  
 CC (allergies), infectious diseases (e.g., viral hepatitis), complement  
 CC activation disorders, immune complex diseases, neoplastic disorders  
 CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders  
 CC associated with neovascularisation, diseases at the cellular level,  
 CC cardiovascular disorders (arrhythmias), wound healing and epithelial  
 CC cell proliferation, endocrine disorders (diabetes mellitus) and  
 CC neurological disorders (ischaemic lesions). Immune-system related protein  
 CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to  
 CC maintain organs before transplantation, to treat weight disorders, to  
 CC modulate mammalian characteristics, to change a mammal's mental or  
 CC physical state, or as a food additive or preservative. Immune-system  
 CC related DNA is useful in gene therapy, for chromosome identification,  
 CC radiation hybrid mapping, long range restriction mapping and in forensic  
 CC biology. The present sequence represents a human immune-system related  
 CC protein of the invention.  
 XX  
 SQ Sequence 204 AA;  
 Query Match 29.2%; Score 1079; DB 22; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-91;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 485 LLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELSAQGPVAFHQRRTLQEGGV 544  
 DB 1 llysaddsgferlvgalasalcolplrvavdlwsrrelsaqgpvafhqr rrtlqeggv 60  
 QY 545 VVLLFSPGAVLCSHWDGVSGPGAGHDPFRASLSCLVLPFLOGRAPGSVVGACFDR 604  
 DB 61 vllfsgpavalcsehwldgvsgpgagghdpfraslscvlpflograpgsyvgacfd 120  
 QY 605 LLHPDAVPALFRTVPVFTLPSOLPDFLQALQQPRAPRSGRLQERAEQVSALQPALDSYF 664  
 DB 121 llhpdavpalftvpvftlpsqlpdfalqqrapsrigrqleraeqvsalqpaldsyf 180  
 QY 665 HPPGTPAPGRGVGPGAGAGDGT 688  
 DB 181 hpgtpapgrgvpgagagdgdt 204  
 RESULT 13  
 AAY76048  
 ID AAY76048 standard; Protein; 617 AA.



```

XX AC AAY76048;
XX XX
XX 27-MAR-2000 (first entry)
XX DE Murine skin cell protein, SEQ ID NO:303.
XX KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
XX KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
XX KW secreted; transmembrane; inflammation; cancer; neurological disease;
XX KW angiogenesis; tumour vascularisation; growth disorder;
XX KW developmental disorder; skin wound; hair follicle disorder;
XX KW anti-inflammatory; cytostatic; neuroprotective; vulnery.
XX OS Mus sp.
XX PN M09955865-A1.
XX PD 04-NOV-1999.
XX PF 29-APR-1999; 99WO-NZ00051.
XX PR 29-APR-1998; 98US-0069726.
XX PR 09-NOV-1998; 98US-0188930.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
XX DR WPI: 2000-072177/06.
XX DR N-PSDB: AAZ61753.
XX PT Novel polynucleotides useful for the treatment of various conditions
XX PT including wounds and cancer -
XX PS Claim 4; Page 179-180; 235pp; English.
XX CC The invention relates to novel nucleic acid sequences derived from rat
XX CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
XX CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
XX CC cells. Polypeptides of the invention may be used to treat inflammation,
XX CC cancer and neurological diseases. The proteins may be used to stimulate
XX CC the growth and motility of keratinocytes, to inhibit the growth of
XX CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
XX CC modulate skin inflammation, to modulate epithelial cell growth and to
XX CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
XX CC to treat growth and developmental defects, skin wounds and hair follicle
XX CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
XX CC by cDNA sequences derived from several mouse, rat or human skin cell
XX CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
XX CC AAY76119 are proteins with an N-terminal signal sequence, indicating
XX CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
XX CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
XX CC putative transmembrane domains.
XX SQ Sequence 617 AA;

```

```

Query Match 8.3%; Score 305.5; DB 21; Length 617;
Best Local Similarity 25.9%; Pred. No. 4.3e-19;
Matches 130; Conservative 53; Mismatches 179; Indels 139; Gaps 18;

```

```

OY 393 STRAARLG-----EYLLDOLSGGCIQLMDDDLGLM---ACP 427
DB 349 awsdpglqdpmpvyslsqtgsvpyldlilpflrqencillwrsdvhfawkhvlp 408
OY 428 MDKYLHKRWALVWLACLFFAALSLILLKKDHAKGWLRLKQDVSGAARGRAALLLY 487
DB 409 dd-----apytqlll-----rslsgtrtpvllln 434
OY 488 SADDGFERLYGALASALCOLPLRYA-----VDMRSRELSAGCPVAMFHAORQTL 539
DB 435 aadsaqrllvgalael-----lrlalgggrivdlwglthvarlqplvwaalerwa 489
OY 540 QEGGVVLLFSPGAVALCSEWLQDGVSGPGANGPHDAFPASLSCVLPDFLGRAPGSYVG 599
DB 490 reggtvlllwn-----cagpstacsgdpqaslrlll-----caaprp11l 530
OY 600 ACFDRLLHPDANPALEFRVYVFTLSQPLDFLGAID-QPRARPSG-----RIQERAE 650
DB 531 ayfstrcakgdlprlralpryllrldlprlraldaqpatlasswshlgakrc1knrl 590
OY 651 QVS-RALQPALDSYFHPGTP 670
DB 591 qchlleaakddygstnsp 611

```

```

RESULT 14
ID AAB55987 standard; Protein: 617 AA.
XX AC AAB55987;
XX DT 08-MAR-2001 (first entry)
XX DE Skin cell protein, SEQ ID NO: 303.
XX KW Mouse; skin cell; cytostatic; anti-inflammatory; anti-HIV;
XX KW neotropic; neuroprotective; vulnery; immunomodulatory; vaccine;
XX KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
XX KW inflammation; neurological disease.
XX OS Mus sp.
XX PN M0200069884-A2.
XX PD 23-NOV-2000.
XX PF 15-MAY-2000; 2000WO-NZ00075.
XX PR 14-MAY-1999; 99US-0312283.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
XX DR WPI: 2001-007495/01.
XX DR N-PSDB: AAC99686.
XX PT New isolated polynucleotide used in the identification of genetic
XX PT disorders and encoding polypeptides used for treating inflammatory
XX PT disease, cancer and neurological diseases -
XX PS Claim 4; Page 244-245; 352pp; English.
XX CC The present sequence is a polypeptide which is expressed in
XX CC mammalian skin cells. The polypeptide is useful for stimulating
XX CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
XX CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
XX CC tumours, modulating skin inflammation, stimulating the growth of
XX CC epithelial cells, inhibiting the binding of human immunodeficiency virus
XX CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
XX CC neurological diseases. The polynucleotide can be used as a marker, in
XX CC the identification of genetic disorders, and for the design of

```

FT  
Region  
104..108  
/note= "cAMP/GMP-dependent protein kinase"

vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, immune-mediated renal disease, a demyelinating disease, an autoimmune or immune-mediated skin disease, contact dermatitis, an allergic disease e.g. food hypersensitivity, asthma, a transplantation associated disease, or a chronic inflammatory demyelinating polyneuropathy. Treating

CC degenerative cartilaginous disorder comprises administering a PRO1031 or  
 CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous  
 CC examples of the diseases and disorders are given in the specification.

XX  
 SQ Sequence 667 AA;

Query Match	7.8%;	Score 287.5;	DB 22;	Length 667;
Best Local Similarity	27.8%;	Pred. No. 2.2e-17;		
Matches 134;	Conservative 54;	Mismatches 175;	Indels 119;	Gaps 22;

  

QY	237	KNLTGPIITLNTHTDLPCLICQVWPLEPDSVRTINCPFPREDPRAH-QNLWQAAAR-----	290
Db	223	kiwsgghtvelpefilpclicaeasylqedtvrkkcpfqswpeaygsdfkshvftdys	282
QY	291	-----LRLTLQSWLLDAPCSLPAEALCWRAFCGDPCCQPLVPPPLSWENVTVDVNSS---	342
Db	283	ghtqmvmtlfr-----cpkleaalqqrhdwtickdl-----pnataresdgwyv	329
QY	343	EKLQIQ-----ECLWADSLGPKLDDVLLLETRGPQ-----DNRSICALE	381
Db	330	lekvdhlpqlcfkfsfgnshvec--phqtgsitwnvsmtdgaqqlilhfssrmhatf-	386
QY	382	PSGCTSLP-----SKASTRAARLGEYLLODLOSGOCLOLWDDDLGALW---	424
Db	387	-saawslpqlgqdtlvppvytvsqargsspsvldlipfipgcccivwrsdqvfwkhl	445
QY	425	ACPMDKYIHKRWALVWLACLLFAAALSLLLLKKHAKGWRLLLKQDVRSGAAARGRAAL	484
Db	446	lcpdvsvyrh-----lglllallalltllgv-----vltctirpqsppg-arpl	491
QY	485	LLYSADDSGFERLVGALASALCOLPLRVA-----VDLWSRRELSAQGPVAFWHAQRR	536
Db	492	lihaadseagrllvgalael-----lraalgggrdvldiwegrhvarvgplwlaart	546
QY	537	QTLOEGGVVLLFSPCAVALCSEWLQDVGSGPGAGHPDAFRASLSCVLPDFLQGRAPGS	596
Db	547	rvareqgtvllwsgadlr-----pvsqp-----dpraaplall-----haaprp	587
QY	597	-YVGACFDRLHPDVPALFRTVPVFTLPSQLPDFLGALQ-QPRAPRS--GRLOERAEQV	652
Db	588	lillayfsrlcakgdipppralprylrldlprllraldarpfaeatswgrlgarrrq	647
QY	653	SR 654	
Db	648	sr 649	

Search completed: May 19, 2002, 13:26:42  
 Job time: 7623 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 15:19:08 ; Search time 115.29 Seconds  
(without alignments)  
1032.358 Million cell updates/sec

Title: US-09-608-918-10  
Perfect score: 3692  
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPRGVGPAGPGAGDGT 688

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP-archaea.\*
- 2: SP-bacteria.\*
- 3: SP-fungi.\*
- 4: SP-human.\*
- 5: SP-invertebrate.\*
- 6: SP-mammal.\*
- 7: SP-mhc.\*
- 8: SP-organelle.\*
- 9: SP-phage.\*
- 10: SP-plant.\*
- 11: SP-rodent.\*
- 12: SP-virus.\*
- 13: SP-vertebrate.\*
- 14: SP-unclassified.\*
- 15: SP-virus.\*
- 16: SP-bacteriaph.\*
- 17: SP-archaeop.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2554.5	69.2	538	4	Q9BR97	Q9br-97 homo sapien
2	1724.5	46.7	567	11	Q99J43	Q99j43 mus musculus
3	176	4.8	866	4	O43844	O43844 homo sapien
4	172	4.7	866	4	O96F46	O96f46 homo sapien
5	166	4.5	864	11	O60943	O60943 mus musculus
6	125.5	3.4	1366	2	Q9Z529	Q9z529 streptomyce
7	120.5	3.3	478	16	Q9RY89	Q9ry89 deinococcus
8	116	3.1	5435	2	Q9L4X2	Q9l4x2 streptomyce
9	115.5	3.1	698	16	Q9PBH5	Q9pbh5 xyella fas
10	115.5	3.1	2424	5	Q9VZ48	Q9vz48 drosophila
11	113	3.1	4928	2	Q9ALM3	Q9alm3 saccharopol
12	112.5	3.0	1193	16	Q9HZE5	Q9hze5 pseudomonas
13	112	3.0	9510	2	Q93NX9	Q93nx9 streptomyce
14	111.5	3.0	679	2	Q9F2P6	Q9f2p6 streptomyce
15	111	3.0	361	2	O86440	O86440 pseudomonas
16	111	3.0	620	4	Q9HCN6	Q9hcn6 homo sapien

17	111	3.0	1118	4	Q15426	Q15426 homo sapien
18	111	3.0	1256	11	Q99M76	Q99m76 rattus norv
19	110.5	3.0	593	2	Q9FB15	Q9fb15 streptomyce
20	110.5	3.0	4340	2	O30764	O30764 streptomyce
21	110	3.0	1318	12	P90493	P90493 herpes simp
22	109.5	3.0	2055	4	O75055	O75055 homo sapien
23	108.5	2.9	712	5	Q9BH06	Q9bh06 leishmania
24	108.5	2.9	6797	2	Q9X993	Q9x993 streptomyce
25	108.5	2.9	10917	2	Q93NW6	Q93nw6 streptomyce
26	107.5	2.9	433	16	Q9RTZ2	Q9rtz2 deinococcus
27	107.5	2.9	604	12	Q9YVB7	Q9yvb7 pseudorabie
28	107	2.9	825	4	Q96P01	Q96p01 homo sapien
29	106.5	2.9	274	16	Q9I615	Q9i615 pseudomonas
30	106.5	2.9	430	4	Q96924	Q96924 homo sapien
31	106.5	2.9	884	2	Q9RJ90	Q9rj90 streptomyce
32	106.5	2.9	912	2	Q9K3Y2	Q9k3y2 streptomyce
33	106.5	2.9	1030	4	Q9BY59	Q9bys9 homo sapien
34	106	2.9	1115	4	Q9HD43	Q9hd43 homo sapien
35	106	2.9	4247	2	Q9L8H4	Q9l8h4 streptomyce
36	105.5	2.9	679	16	Q9REE6	Q9rre6 deinococcus
37	105.5	2.9	1060	4	Q9PI27	Q9pi27 homo sapien
38	105	2.8	757	6	Q9N015	Q9n015 macaca fasc
39	105	2.8	766	11	Q9JKQ9	Q9jkq9 mus musculus
40	105	2.8	785	2	Q9K3N5	Q9k3n5 streptomyce
41	105	2.8	921	4	Q9UJ19	Q9uj19 homo sapien
42	105	2.8	3670	2	Q9Z4X5	Q9z4x5 streptomyce
43	104.5	2.8	238	16	O53500	O53500 mycobacteri
44	104.5	2.8	1343	12	Q06635	Q06635 bovine herp
45	104.5	2.8	1505	2	Q9F0D7	Q9f0d7 streptomyce

#### ALIGNMENTS

RESULT 1

ID Q9BR97 PRELIMINARY; PRT; 538 AA.

AC Q9BR97;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 59.1 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC006411; AAH06411.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;

Query Match 69.2%; Score 2554.5; DB 4; Length 538;  
Best Local Similarity 96.5%; Pred No. 6e-200;  
Matches 475; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

Qy 1 MPVPWFLLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLWSDILCLPGDIPVAPGPV 60

Db 1 MPVPWFLLSLALGRSPVLSLERLVGPQDATHCSPGLSCRLWSDILCLPGDIPVAPGPV 60

Qy 61 LAPTHLOTELVRCKETDCDCLRLRVAVHLVGHWEPEDEEFKFGGAADSGVEPRNAS 120

Db 61 LAPTHLOTELVRCKETDCDCLRLRVAVHLVGHWEPEDEEFKFGGAADSGVEPRNAS 120

Qy 121 LQAQVLSFQAYPTARCVLLEVQVPAALVQFGSGVGVVYDCFEALGSEVRWISYTOPR 180

Db 121 LQAQVLSFQAYPTARCVLLEVQVPAALVQFGSGVGVVYDCFEALGSEVRWISYTOPR 180

Qy 181 YEKELNHTQOLPALPWLNVNSADGNVHLVNVSEHQHFLGLSLYWNQVQPPKPRWHKHLT 240

```
Db 181 YEKELNHTQOLPALPWLNVNSADGNVHLVNVSEEHGFLSLVNWQVGGPKPRWHKNT 240
Qy 241 GQIITITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLQSWL 300
Db 241 GQIITITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLQSWL 300
Qy 301 LDAPCSLPAEALCWAPGDCQCPVLPPLSWENVTVVD-----VNSSE 343
Db 301 LDAPCSLPAEALCWAPGDCQCPVLPPLSWENVTVVDKVLFPKLLKHGHNLCVQVNSSE 360
Qy 344 KLQLOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 403
Db 361 KLQLOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEYL 420
Qy 404 LDQLOSGQCQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALSLILLKKDHAKG 463
Db 421 LDQLOSGQCQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALSLILLKKDHAKG 480
Qy 464 WLRLKQDVRS 475
Db 481 WLRLKQDVRS 492

RESULT 2
Q99J43 ID Q99J43 PRELIMINARY; PRT; 567 AA.
AC Q99J43;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 62.8 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004759; AAH04759.1;
KW Hypothetical protein.
SQ SEQUENCE 567 AA; 62798 MW; CLAAAB79E2006B1D CRC64;
```

```
Query Match 46.7%; Score 1724.5; DB 11; Length 567;
Best Local Similarity 59.9%; Pred. No. 3e-132;
Matches 348; Conservative 47; Mismatches 125; Indels 61; Gaps 9;
```

```
Qy 1 MPVFWFLSLALGRSPVLSLERLVGPQDATHCSGLSCLRLWDSIILCLPGDIVPAPGPV 60
Db 1 MPVFWFLSLALGRNPVVSLSERLMEPQDTARCISGLSCHLWDGDLVCLPGLSQAAPGPV 60
Qy 61 LAPTHLOTELVLRQKETCDCLRLVAVHLAVHGHWEPEDEKFGGAADSGVEEPNAS 120
Db 61 LVPTRLQTELVRQCPQKTCALCVRVVHLAVHGHWAPEE-----AGKSDSELOESNAS 116
Qy 121 LQAQVYLSFOAYPTARCVLLEVOVPAALVQFGOSGVVYDCFEAALGSEVRINWSTQPR 180
Db 117 LQAQVYLSFOAYPTARCALLEVOVPAALVQFGOSGVSAVDFCEAALGSEVRINWSTQPR 176
Qy 181 YEKELNHTQOLP-----ALPWLNVSDGDNVHLVNVSEEHGFLSLVYN 225
Db 177 YQKELNLTQOLPDCRGLVDRDSTQSCWVLPWLNVSTGDGNNVLTLDVSEQDFSLYLRL 236
Qy 226 QVGGPKPRWHKNTGQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHONL 285
Db 237 PVPDALKSLWYKNTLGPQITLNTDLVPCICIQWNSLEPDSERVEPCPFREDPGAHRNL 296
Qy 286 WQAARLLTLQSWLIDAPCSLPAEALCWAPGDCQCPVLPPLSWENVTVVD----- 338
Db 297 WHIARLRVLSPGVWQIDAPCCCLPGKVTLCWQAPDQSPQPLVPVPVPOKNATVNEPQDFQL 356
```

```
Qy 339 -----VNSSEKLOIQLQOECLWADSLGPKDDVLLLETRGPQDNRSICALPESGCTSL 388
Db 357 VAGHPNLNCVQYSTWEKVLQACWLADSLGPFKDDMLLVEMKTGLNNTSYCALEPSGCTPL 416
Qy 389 PSKASTRAARLGEYLQLOSGQCQLW- DDDLGALWACPMCKYTHKRWALVWLACLLFA 447
Db 417 PSMASTRAARLGEELLQDFRSHOCQMLWDDNMGSWACPMCKYTHKRWALVWLACLLFA 476
Qy 448 AALSILLILLKKDHAKGWLRLKQDVRSAAAARGRAALLYS-----ADDSGFERLVGAL 501
Db 477 AALFEFLLKKD-----RMTSPSPGSAACYPIACKAGRPAAATSG-STSTGCC 522
Qy 502 ASALCQPLRVAVDLWMSRRELSAOGPVAFWHAQRQTLOEG 542
Db 523 QTTLCT--PPRSASPRSSPCPRSCR--LSWMHCREAAPLPRG 559

RESULT 3
O43844 ID O43844 PRELIMINARY; PRT; 866 AA.
AC O43844;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035683; PubMed=9367539;
RA Yao Z., Springs M.K., Derry J.M.J., Strockbine L., Park L.S.,
RA VandenBos T., Zappone J., Painter S.L., Armitage R.J.;
RT "Molecular characterization of the human interleukin (IL)-17
RT receptor.";
RL Cytokine 9:794-800(1997).
DR EMBL; U58917; AAB99730.1;
KW Receptor.
SQ SEQUENCE 866 AA; 96122 MW; 88AF626A83F3FF70 CRC64;
```

```
Query Match 4.8%; Score 176; DB 4; Length 866;
Best Local Similarity 21.8%; Pred. No. 9.7e-06;
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;
```

```
Qy 237 KNLT--GQIITLNL-----HT---DLVPCICIQWPLEPDS-----VRTN-- 271
Db 66 RNLTSPSPKDLQIQLHFAHTQGGDLFPVAHIE-WTLQTDASILYLEGAELSVLQLTNR 124
Qy 272 IC---PFREDPRAHONLWQAARLLTLQSWLIDAPCS-----LPAAALCWAPGDC 321
Db 125 LCVRFEEFLSKLRHHRRNR-----FTFSHFVVDQVEYEVTVHLPKP-----IPGD 172
Qy 322 PCQP-----LVPLPS-----WE-NYTVDNVNSSEKLOECLWADSLGPIK 360
Db 173 PNHQSKNLFVPCDEHARKVVTTPCMSGSLWDPNITVTEAHQLRVSTFWNES----- 227
Qy 361 DDVLLLETRGPQDNRSICALPESGCTSLPSKASTRAARLGEY-----LLQIDQ 408
Db 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEFHORSNVTTLRLNK 274
Qy 409 S-----GQCILWDDDL--GALWACP-----MDXYIHKRWALVWLA 442
Db 275 GCCRHQVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPIPDYMPLMWV--WFTIGIS 329
Qy 443 CLLFAAALSLILLKKDHAKGWLRLKQDVR--SGAAA-----RGAALLLSADDS 492
Db 330 ILLVGSVILLIVCWTRWRLAGPSEKYSDDTKYTDGLPADLIPPLPKRWIIYSADHP 389
Qy 493 GFERLVGALAS--ALCQPLRVAVDLWMSRRELSAOGPVAFWHAQRQTLOEGGVVLLFSP 551
Db 390 LYVDVWLKFAQLLTACGTEVALDILEQATSEAGVMTWVGKQKQEMVSNESKIIIVCSR 449
```

Db	505	SEVSGDGVDPDLFGAAPRYPLMDRFEEVYFRIQDLEMPQPGRM-HRVGELSG	-----DN 557
Qy	663	YFHPPG 668	
Db	558	YLRSPG 563	
RESULT	5		
ID	Q60943	PRELIMINARY;	PRT; 864 AA.
AC	Q60943;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	INTERLEUKIN 17 RECEPTOR.		
GN	IL17R.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=THYMOMA EL4;		
RC	MEDLINE=36111968; PubMed=87777726;		
RA	Yao Z., Fanslow W.C., Seidin M.F., Rousseau A.M., Painter S.L.,		
RA	Comeau M.R., Cohen J.I., Spriggs M.K.;		
RT	"Herpesvirus Salmiri encodes a new cytokine, IL-17, which binds to a		
RT	novel cytokine receptor."		
RL	Immunity 3:811-821(1995).		
DR	EMBL; U31993; AAC53557.1; --		
DR	MGD; MGI:107399; I117r.		
SK	Receptor.		
QW	SEQUENCE 864 AA; 97807 MW; 343FD51AA687DA31 CRC64;		

Query Match	4.5%	Score 166;	DB 11;	Length 864;
Best Local Similarity	20.4%;	Pred. No. 6.3e-05;		
Matches 126;	Conservative 77;	Mismatches 196;	Indels 226;	Gaps
QY 192	PALPW----	LNVSADGNVHLVLN----	VSEOHFGLSLYWNVOGPP--	KPRW--HKNL 239
DB 14	PALGWLLLLLVN	LWVGRASPLDLFPAPVCAQE--	GLSC---	RKVNSTCLDDSHIHPKNL 68
QY 240	T--GPQIITLN----	HTDLPCLCIQVWPLEPDS----		VRTN--IC- 273
DB 69	TPSPKNIYINLS	VSSYQHGEVPLVHVE-WT	LQTDASILYLEGAELSVLQNTNERLCV	127
QY 274	--PREDPRAHQN	LWQAARLLTLQSWLLDAPCSLPAEA-----	ALCWRA	GGDP--- 322
DB 128	KFQFLSLQLHHR	KWR-----	FSFSHFVVD-----	PGQEYEVTVHHLKRPIDPGDPNHR 176
QY 323	-----	CQPLVPPLS-----	WE-NVTVDVNSSEKLOLOECLWA	DLSGLPKLDDVL 364
DB 177	SKIIIVPDCED	CKMKMTSCVSSGSLMDPNITVETL	QHLRVDFLTWNST-PYQ--	V 232
QY 365	LLETRGPDNRSL	CALEPSPGCTSL-----	PSKASTRAARLGEVLLQDLQSGCQLWDD	418
DB 233	LLESFSDSENHS-----	CFDVVKQIAPAPQEEFHQRANVT	FLSKFH-----	274
QY 419	DLGALWAC-----			PMDKYIHKRWA--LV 439
DB 275	-----	WCCHHHVQVOPFFSSCLNDCLR	HAVTVPCPVISNTTVPKPADYI-PLWVYGLIT	328
QY 440	WLACLLFAAALS	LTLLKKDHAGWRLRLKQDVRSGA-----		476
DB 329	LIAILLGVSVTL	IICHT-----	W-RL-----	SGAQEKHGDDSKINGILPVALTTP 375
QY 477	AARGRAALLYS	ADDSGFERLVGALASAL-CQLPLRV	AVDLWSRREL	SAGQPVAFWFAOR 535
DB 376	PLRPKRWIVTS	ADHPDLYVEVWLKFAOFLIT	ACGTETVALDLLEE	OVISEVGVMTWVSROK 435
QY 536	RTQLQEGGVVLL	FSPG-----		AVAL-CSEWLODCVSGPAHGHPDAFRAS 580

```
Db 436 QEWESNSKIIILCSRGTOAKWKAILGWAEPVQLRCDHWKPG- - - - -DLFTAA 485
QY 581 LSCVLPDFLOGRAPSGYSGACFDRLHDPDAVPALFTVPVFTLPSPQLPDFLGLAQPPAP 640
Db 486 MNMILPDFKRKPACFGYVVCYFSGICSDRDVPDLFNITSRYPLMDREVEEYFIQDLEMF 545
QY 641 RSGRLQERAEQVSRLQALPDLSYFHP 667
Db 546 EPRMHVRELATG- - - - -DNYLQSP 565

RESULT 6
Q92529 ID Q92529 PRELIMINARY; PRT: 1366 AA.
AC Q92529;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PUTATIVE LARGE PRO/ALU/GLY-RICH PROTEIN.
GN SC9F2.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035559; CAB37473.1; -
SQ SEQUENCE 1366 AA; 141042 MW; FB954569D87D029B CRC64;

Query Match 3.4%; Score 125.5; DB 2; Length 1366;
Best Local Similarity 23.2%; Pred. No. 0.23;
Matches 177; Conservative 58; Mismatches 248; Indels 281; Gaps 43;

QY 21 LERLVGPQDATHCSGLSLRLWSDILCLPGDITVAPGPVLIATHQTELVLRCQKETDC 80
Db 675 LDGALADPDAGH- - - - -PLTIRLLSEVRAALFGP- - - - -PAPVPVTRDAVFTAYLDL- - - - - 721
QY 81 DLCLRVAVHVLAVGHWEEDKEKFGAA- - - - -DSGVEPRNASL-QAOWVLSFOA 131
Db 722 -MCLRVATRLA- - - - -DENLGRGTAVRLAAKVSQGVHEAARRSLGPGGGLDRES 771
QY 132 YPT- - - - -ARCVLLE-VQVPAAL- - - - -VQFGQSVG 156
Db 772 FETLFPCCPAPARLGGTGWAPAVLAEGLVPTGSGYRFAHEELADWIOGTHIDLGEALR 831
QY 157 SVVYDCFEAALGSEVRINSYTOPRYEKELNHTQOLPALPMLNYSADGDNVHLVNLVSEEQ 216
Db 832 ALVHR-RDTPLGT- - - - -HTRTL-VPVHRI- - - - -GSVEALLLLA-RQ 868
QY 217 HGLSLYNQVQGPVPRVHKNLTGQITLNLHTDLVPLCICIQVWPLEPDSVRTNCPFR 276
Db 869 H- - - - -GVPO- - - - -LALTLEELVHAL- - - - -D 886
QY 277 EDPRAHQLNWAARLRLLTLQSWLLDAPCSLPAEAL- - - - -CWRAPGDPQCLVPP 329
```

```
Db 887 RDPHS- - - - -WAAKL- - - - -LAELTRVPDATPYTDVLLADGIAERAGDQPTQVFGP 938
QY 330 LSWENVTVDVNSSEKQLQECLEWADSLGPKDDVLLLETRGPQDNRSLEALEPSGCTSLP 389
Db 939 AFWTAPRVPATRLDL-LRLVLAD- - - - -GPPH- - - - -EPG- - - - -P 970
QY 390 SKASTRAARLGEYLLQDLSQGCQLQ-LWDDDLGALWACPMQKYIHKRWALVWLACLLFAA 448
Db 971 RHLDTAAG- - - - -LLVADPRTVQPLLVRFDDERPLPATP- - - - -HATVA- - - - -TA 1012
QY 449 ALSILLILLKKDHAKGWLRLIKQDVRSAAARGRAALLLYSADDSFERLVGALA- - - - -SA 504
Db 1013 AQUALTHRHRLGDLGLTEVLVDSTHRA- - - - -DELLAVLAEEPSA 1054
QY 505 LCQLPLRVAVDLWSRRLSQAQPVAVFHAQR- - - - -QTLOEGGVVLLFSP- - - - - 551
Db 1055 LCR- - - - -AVERWARDERPAHRAAVTHGLTAPHARGADFTLLRHAALVLLAGPSDP 1109
QY 552 - - - - -GAVALCSEWLQDVGSGAGHPH- - - - -DAFRA- - - - -SLSCVLPDFLOGRAPG 595
Db 1110 LRGGALALL- - - - -VQDPDRHLPAALDLFAACDPLPPSAVAAALPTH- - - - -PE 1157
QY 596 SYVCACFDRLHDPDAVPALFR- - - - -TVPVFTLPSPQLPDFLGLAQOPRPRSRLOERAQ 651
Db 1158 PVLEAFRALILGPDAGEALRLADATTPLT- - - - -HRVAALVGRVTVPETAGHL- - - - -AAY 1212
QY 652 VSRL- - - - -OPALDSYFHPPT- - - - -PAPRGVGVGAGPGAGDG 687
Db 1213 VDRLDRDPAPRAVLLPLVTRLLDDGPEPARAA- - - - -LAGVLAADG 1254

RESULT 7
Q9RY89 ID Q9RY89 PRELIMINARY; PRT: 478 AA.
AC Q9RY89;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HYPOTHETICAL 51.2 KDA PROTEIN.
GN DR0061.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AF001869; AAF09658.1; -
DR TIGR; DR0061; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 478 AA; 51231 MW; 5E74CDEB8FBA5C95 CRC64;

Query Match 3.3%; Score 120.5; DB 16; Length 478;
Best Local Similarity 24.3%; Pred. No. 0.15;
Matches 99; Conservative 39; Mismatches 148; Indels 121; Gaps 17;

QY 276 REDPRAHQLNWAARLRLLTLQSWLL- - - - -DAPCSLPAEALCWRAPGDPQCLV 327
Db 110 RADPRERPLLEAAQ- - - - -RIRRLARLARAPGLPTVPAHSVLY- - - - -AQEASALS 161
QY 328 PPLSWENVTVDVNSSEKILQLECL- - - - -WADSLGPKDDVLLLETRGPQDNRSLEALEPSG 384
```

```
Db 162 APAAGDGLTHLGGPDRLAHQGLRLSWR-SLGPWNQQLLVQDETSPQVSGGQALLRP-- 218
QY 385 CTSLPKASTRAARLGEYLLDQSGQCQLWDDDLGALWACPMCKYIHKRWALVWLACL 444
Db 219 --DUPASERQLFTVGGQQLVLFSG----- 242
QY 445 LFAAALSILILLKXDKHAKGWLRLKQDVRSGAAARGRA-ALLLYSADDSGFFERLYGALAS 503
Db 243 -----DYVLLRRRADAQTHLAR-----LAALGRACALLLLPAEQHGRRLRLARTLAR 290
QY 504 ALCOLPLRVAVDLWSRRELSAQQVAVWFHQAORROTLOEGGVVLLFSPGAV----- 554
Db 291 RLRGDPPR-ADDPASGQVAFATFAECLEAARR-SLQRLEPLILGRFSPAQAHAHEAQAQA 348
QY 555 -----ALCSEWLQDVSGP--GAHGPHDAFRASLSCVL-----PDFL 589
Db 349 LLHLPPHYAAQVQAAEHAASAEPLREAADTDLVGREANGPVLPATGHFIVFLPGPDPL 408
QY 590 QGRAPGSVVGACFDRLL--LHPDPAVPALEPRTVPVFTLPOLPDLFGLAL 634
Db 409 EVRLPG-----DRLLTLRPD-----YRAELVAVLPQAAVTVGDL 443
```

## RESULT 8

```
Q9L4X2 ID Q9L4X2 PRELIMINARY; PRT; 5435 AA.
AC Q9L4X2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NYSJ.
GN NYSJ.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1971;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11455;
RX MEDLINE=20334850; PubMed=10873841;
RA Brantaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
RA Valla S., Zolchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway.";
RL Chem. Biol. 7:395-403(2000).
DR EMBL; AF263912; AA071767.1;
DR HSP; P25715; IMLA.
DR InterPro; IPR001227; Acyltransf_domain.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Phosphopant_attach.
DR Pfam; PF00698; adh_trans; 3.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 5435 AA; 562659 MW; AA55465DF087A38C CRC64;
```

Query Match 3.18; Score 116; DB 2; Length 5435;  
Best Local Similarity 24.88; Pred. No. 8.3;  
Matches 118; Conservative 44; Mismatches 196; Indels 118; Gaps 27;

```
QY 290 RLRLTLTQSWLLDAPCSLPAAALCWAPGDPCCOPLVPPLSWENVTVDVNSEKILQOE 349
Db 2817 RVRRLAV-SHAFHSPMDPMLDAFRAVAEGLEYHQPRIPVVS--NVTGEVAAAEELCAAD 2873
```

```
QY 350 CLWADSLGP---LKDDVLLLETRCPQDNRSICALPESGCTS-----LPSKASTRAARLG 400
Db 2874 -YVVRHVRAVRFPADGVRTLAERGA---TAFLEIGPDGVLSALARGVLPAAELVPT-- 2926
QY 401 EYLLQDLQSQCLQWLDDDLGALWACPMCKYIHKRWALV-WLACLFLFAAALSLLFL-LKK 458
Db 2927 --LRKORDEESAL-----LAGL-----ARLHVAGVTVDWSAALTGTGARGTDLPTIYAF 2972
QY 459 DHAKGWLRLKQDVRSGAA-----RGRAALLLYSADDSGFFERLYGALASALCOL 508
Db 2973 QREYWPPELAEPAGGADAADAEFAAVERADATATAAHLIDGDO--LGAVLPALSAW 3030
QY 509 PLRVAVDLWSRRELSAQQPV---AMFHAORROTLOEGGVVLLFSPCAVALCSEWLQDG 564
Db 3031 RTR-----RRTTSATNALRHRESWEPLSLAGTPHTGGVLVLV---PAAATDPPWADV 3080
QY 565 VS--GPCAHG---PHDAF-RASLSCVLPDFLOGRPGSVYG-ACFDRLLHPDAVP----- 612
Db 3081 VAALGPDARRVDPADGTDRALAAALTEAADDTAPTAVVSLALDETSGDDAVPAGTTA 3140
QY 613 -----ALFRT---VPVFTLP-----SOLPDFLGAQOPRAPRS-----GRL----- 645
Db 3141 TAALVQALADTGAPLWALTRGAVPALPD-----EQTAPAAQAVWGLGRIALALELPRH 3195
QY 646 -----QERAEQSVSRALQPAL-----DSYFHPGPTAPGPGVGPAGPGAGDGT 688
Db 3196 WGLGLVDLPADLDETRARLLPAALADAGDEQDLALRATGAYGRRITPAPADDPAGT 3251
```

## RESULT 9

```
Q9PBH5 ID Q9PBH5 PRELIMINARY; PRT; 698 AA.
AC Q9PBH5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF2169.
GN XF2169.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Narino C.L.,
RA Marches M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
```



RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL: AE004030; AAF84968.1; .  
DR InterPro: IPR001440; TPR.  
DR Pfam: PF00515; TPR; 6.  
DR SMART: SM00028; TPR; 3.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 698 AA; 77588 MW; 91326560086F5A7F CRC64;

Query Match 3.1%; Score 115.5; DB 16; Length 698;  
Best Local Similarity 21.8%; Pred. No. 0.63;  
Matches 160; Conservative 90; Mismatches 234; Indels 251; Gaps 43;  
Qy 61 LAP-----THLQTVLVRCKE--TDCDLCLRVAVHLAVHWEPEDEKPG-----G 107  
Db 69 LAPERPELHLQGLLHAQHDLSGDAALSRL-----DPNQFTYVMOAHLA 118  
Qy 108 AADSGVEPRNASLQAVVLSFQAVP-----TARCVLLEVVQVPAALVQGSV-----G 156  
Db 119 IARGALDAEQLSRTRATRIIS--DHPQLLVNATIALRGQVDOALTMLSQAVELLPEEP 176  
Qy 157 SVVYDCFEAALGSVRINWYTPRY-----EK-----ELN-HTQQQLPALPWLNVASDGDN 205  
Db 177 AVLE-----SLG-----FAYLQKHGHTAFERAFAFORVIELNPHVIPRYALIAQLAQRG-R 225  
Qy 206 VHLVNLVSEEHFGLSLYVNVQVGPCK-----PRWKNLTGPQITLNTDLVPCLCIQV-- 260  
Db 226 LDDALRILEE-----ALSL-----PEGDTPALHR-LTGEFELLAGHPDRALTHLRQVLA 273  
Qy 261 -WPLEPDSVRNINCFRE-----DPRAHNLWQAARLR----- 293  
Db 274 TWPEDPRLEALLTAWKOLDMDDDARITLDAALDIKPRNH-DLW-LARLAVAPVGSDEAR 331  
Qy 294 LTQSLLDAPCSLPAEAALCWRAFGDPCQPLVPPLSVENVTVDVNSSEKLQLECLWA 353  
Db 332 IVIERWLSAMPEHLPALETLM----- 352  
Qy 354 DSLGLKDDVLLLETRGPQDNRSICALEPSGCTSLFSKASTRAARLGEVILQ--DLQSGQ 411  
Db 353 -SLHDIQNNPEAAETVA---RQIVALEPGRLSG-----EQRIVEALLQRDPPTAVA 399  
Qy 412 CLQLWDDDLGALWACPMKDIHKRWAL-VWLACL-----LFAAALSILLLKKDHAKGWL 465  
Db 400 CLQQLIESVPE-----HERITLREWGLVQDRAGQFEAALATWLFQFAEQAKYRL 449  
Qy 466 RLLKQ-----DVRGGAARGAALLLYSADDSGFERLVGALA---SALC----- 506  
Db 450 PLPQTWTNSKQWPDIAIPAEILARPLLLINGPGGSHVERTVMMWASMLCADRYSET 509  
Qy 507 -----QLPLRVAVDLMSRRELSAQGP-----VAWFH-----AQRR 536  
Db 510 PPADPLQRYETVSELTSGTLTPQALIDAW-RQQLPARGIEDGNVIDWLLWDSNLSLTALR 568  
Qy 537 QTLQEGGVVLLFSPGAVALSSEWLQDGVSGPAGHPDAFRASLSCVLPDFLQGRAPGS 596  
Db 569 PHLPEGRLLIIVLRDPRMLL--DWIAYGSPILP---LDSLQQAANW-LGDIL-----N 616  
Qy 597 YVGACFDRLHP-----DAYPALFRTV-PVETLPSQLPDFLQALQOQPRAPRSRL 645  
Db 617 QIAALHEDLPHHILRIDGIEDNPQALATTLEDIFGSPFPIPP-----SLEAPRLP-AGRW 672  
Qy 646 QERAEQVSRALQPAL 660  
Db 673 RDYREVLSAFAVIL 687

RESULT 10

Q9VZ48

ID Q9VZ48

AC Q9VZ48

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

Query Match

Best Local Similarity 3.1%; Score 115.5; DB 5; Length 2424;

Matches 128; Conservative 33; Mismatches 169; Indels 175; Gaps 27;

Qy 269 RTNCPFPREDPRAHQNLWQAA-RLR-----LTLQSWLLDAPCSLPAEAALC----- 314

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE CG2174 PROTEIN.

GN CG2174.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.C., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.G., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saudefs R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RL "The genome sequence of Drosophila melanogaster.";

DR EMBL: AE003484; AAP47980.1; .

DR HSSP: P08799; 1MND.

DR FlyBase: FBgn0030252; CG2174.

DR InterPro: IPR000048; IQ.

DR InterPro: IPR001609; myosin\_head.

DR InterPro: IPR000857; MyTH4.

DR Pfam: PF00612; IQ; 3.

DR Pfam: PF00063; myosin\_head; 1.

DR Pfam: PF00784; MyTH4; 1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR PRODOM: PD000355; myosin\_head; 1.

DR SMART: SM00015; IQ; 3.

DR SMART: SM00242; MYSC; 1.

DR SMART: SM00139; MyTH4; 1.

SQ SEQUENCE 2424 AA; 267616 MW; 8AD62AA33F9AA5D0 CRC64;

Db 906 RRRVQOLRSDHRRRQREAREAREAKQAVLERSQLSYLD----IPAELAFIYSKL 961  
 QY 315 --WRAPGDCPCQPLVPPLSWENTVDVNSSEKIQLOECIMWADSLGPKD---DVLLETR 369  
 Db 962 QGWSPPHGD--RHLVRVLG-----TVPGPPSAVOLPEDLGQFSFGKSSVYCNGLRLQPR 1015  
 QY 370 GPQDNRESLCALEPSGCTSLPSKASTRAARLGEVLLQDLQSGQCQLWDDDL---GALWAC 426  
 Db 1016 -----REP-----ITAPLLTRAASRDQDFQDALAVFKLRLWSNDKALEGAKEKL 1060  
 QY 427 PMDKYIHKR-----WALV--WLACLLFAAALSILI 453  
 Db 1061 LADYIVHKALSSRGLRDEILVOLCNQVHGLPPNSGEATRLWQLGQCCLCFQPSAASFY 1120  
 QY 454 LL-LKKDHAKGWLRL--LLKODVR-----SGAAARGRA-----ALLYLSADDSGFERLV 498  
 Db 1121 LMRFVDEAPESLRPLLLRQLLQOQGGTSSGAVGAGACRSFVPAWLEWRAWTGCD--- 1177  
 QY 499 GALASALCOLPL-----RVAVDLWSRRELSAQGPVAFHQAQRRTLOEGGVVLLFS 550  
 Db 1178 -----MALPLTLPDEASQTVAVDSWTSCEAAALAVSSLGVASR-----GWLTVL--- 1222  
 QY 551 PGAVALCSEWLQGVSGGAGHGFDAFRASLSCVL-----PDFLQGRAPGSY 597  
 Db 1223 -----DQOQLTDSGLDYVMDLIAEKLCLCPAFPAPRSDLLR 1259  
 QY 598 VGACFERLLHPDAV--PALFTVTVFTLPSQLPDFLQALQOOPRAPRSRGLQRAEQVSRA 655  
 Db 1260 SGAKFARTLPDAVKRPA---VP-----PPAPPTSSGKEDVPRERRSR-----ELLSRS 1306  
 QY 656 LQPAL-DSYFHPHGGTAPGRGVCPG 679  
 Db 1307 --SALNERYFEREPSPGQGSSTG 1329

## RESULT 11

QYALM3 ID Q9ALM3 PRELIMINARY; PRT; 4928 AA.  
 AC Q9ALM3;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE POLYKETIDE SYNTHASE EXTENDER MODULES 5-7.  
 GN SPND.  
 OS Saccharopolyspora spinosa.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Pseudonocardineae; Pseudonocardiaaceae;  
 OC Saccharopolyspora.  
 OX NCBI\_TaxID=60894;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21257765; PubMed=11358695;  
 RA Waldron C., Matsushima P., Rosteck P. R. Jr., Broughton M.C.,  
 RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;  
 RT "Cloning and analysis of the spinosad biosynthetic cluster of  
 Saccharopolyspora spinosa."  
 RL Chem. Biol. 8:487-499(2001).  
 DR EMBL; AY007564; AAG23263.1; .  
 DR HSSP; P25715; 1MLA.  
 DR InterPro; IPR001227; Acyltransf\_domain.  
 DR InterPro; IPR000794; ketoacyl-synt.  
 DR InterPro; IPR003880; Phosphopant\_attach.  
 DR Pfam; PF00698; Acyl.transf. 3.  
 DR Pfam; PF00109; ketoacyl-synt. 3.  
 DR Pfam; PF02801; ketoacyl-synt\_C; 3.  
 DR PROSITE; PS50075; ACP\_DOMAIN; 3.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 3.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.  
 KW Phosphopantetheine; Transferase.  
 SQ SEQUENCE 4928 AA; 517247 MW; C22C6361888F865A CRC64;

Query Match

3.1%; Score 113; DB 2; Length 4928;

Best Local Similarity 21.7%; Pred. No. 13;  
 Matches 147; Conservative 66; Mismatches 238; Indels 226; Gaps 31;  
 QY 104 RFG--GAADSGVEPRNASLQAVVLSFOAYPTARCVLLEVOVP-AALVQFGQSGSVY 160  
 Db 4128 EFGPDGALSALVEQCLAGSDQAGRV--AAIPLMRRDRDEVTAVAAALAHVHRGAVDM 4184  
 QY 161 DCFEALAGS---EVRIWSTQPRYEKELNHTOOLPALPMLNVSADGDNVHLVNLVSEQH 217  
 Db 4185 SACEAGTGARTVELPTAFQORY-----WLAGQADGCGDVVADPVDAR- 4229  
 QY 218 FGLSLYWNQVQ--GPKPRWHKNIQTGPIITLNHTDLVPCLCIQVWPLEPDSVYTNICPR 276  
 Db 4230 -----FWELVERADPEP-----LVDELCL-----DRDQ-----PFR 4255  
 QY 277 E-----DPRAHONLWO-----AARLRLLTLOSLLDAPCSLPAEAA 312  
 Db 4256 EVLPVLASMRKQROEALADSWRYQVRWSVEVPSAALRGV---WLVLVLPADVPRD-- 4309  
 QY 313 LWRAPGGDPCQP--LVPLSWENTVDVNSSEKIQLOECIMWADSLGPKDVLLETRG 370  
 Db 4310 -----QPAVVIDALARGAEVAVLELQDQORSALVDKVRVIAD--RTEVTG 4356  
 QY 371 PQDNRESLCALEPSGCTSLPSKASTRAARLGEVLLQDL-QSGQCQLWDDDLGALWACPM 429  
 Db 4357 V---LSLLAMDGMPCAAHPHLSRGVAATV--ILTQVLGDAGVSAPLWLATTGGVEAGTD 4411  
 QY 430 KYIHKRWALVWLACLLFAAALSILI LLLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSA 489  
 Db 4412 GPADPDHGLIW-----GLGRVVGLEHPQWNGGLID-----LPETL 4446  
 QY 490 DSGFERLVGALASALCOLPLRV-AVDLWSRREL-SAQGPVAFHQAQRRTLOEGGVVVL 547  
 Db 4447 DETSRNGLVAALAGTAAEDQLAVRSSGLFVRVRAARNP-----RSETWRSRGTVLI 4499  
 QY 548 LFSPGAV-ALCSEWL-----QDGVSGPGAGH-----PH 574  
 Db 4500 TGGTGALGAEGAVARWLARRGAELHVLISRRGPEAPGAADLGAELTELGVKVTVLACDVTD 4559  
 QY 575 DAFTASLSGV-----LPDFLQGRAPGSYVGACFDRL 606  
 AC Q9HZE5; PRELIMINARY; PRT; 1193 AA.  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA3063.  
 GN PA3063.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,



```
RN RP SEQUENCE FROM N.A.
RC STRAIN=97000351; PubMed-8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL442120; CAC09545.1; -
DR InterPro; IPR002296; N12N6_mtfase.
DR PRINTS; PR00507; N12N6MTFRASE.
KW Hypothetical protein.
SQ SEQUENCE 679 AA; 72161 MW; 514562165037F451 CRC64;

Query Match 3.0%; Score 111.5; DB 2; Length 679;
Best Local Similarity 23.8%; Pred. No. 1.3;
Matches 120; Conservative 43; Mismatches 177; Indels 165; Gaps 30;

Qy 294 LTQSLWLDAPCSLPAEALCWAPGDCQPLVP-----PLSWENVTVDV 339
Db 56 LRAQKLAIEVP---PRERV--WQLAGHPGPVAAALVQAGCVLLIHDRPTLW--LDASA 108
Qy 340 NSSEKLQLOECLWADSLGPKDDVLLLETR---GPQDNRSCLALEPSGCTSLFSKASTRA 396
Db 109 GSDERL-----ADLLPPALDOV--LDARFGTGPQ---RAVTTAGPRLPLPSAPLLRG 155
Qy 397 AR-----LGEYLQDLSQGLQLQDWD--DLGALWACPMCKYTHKRWALV 439
Db 156 LAELAGPGARKAYEFLLGRHL--DANPRQYTLTPDPLADLMAELAGPARTVLDP----- 208
Qy 440 WLAC-----LLFAAALS-----ILLKKDHAKGWLRLKQDVRSGAAAGRAALLYSAD 490
Db 209 ---ACGTGSLRAAAATTPGQELYGOESDPAALATLRLALSTDAVR-----IAAGD 260
Qy 491 DSGFERLVGALA-SALCQLPLRVAVDLWSRRELS-----AQGPVAMP-HAQR 536
Db 261 SLRADARTGLRADAALCHPPFNER--NWGHDELAYDPWEYGPFPARTSELAWQHALAR 318
Qy 537 QTLQGGVVVLLFSCAVA-----LCSEWLQDG-----VSGP-GAHPHDAFRASLSCV 584
Db 319 ---VRDGGTVVLLMPAAASRRSGRRVRADLRRLRGALHVAIALPVGAAPPYN----- 367
Qy 585 LPDFL-----QGRAP-----GSYVG-----ACFDRLLHPDA 610
Db 368 LPLHLWLRPRERAPQPCVLLADTGQFAGEGGGPDWRSVRDAVLDATFTDRTGRLDD 427
Qy 611 VPALFRTVPVFTLPSPDLGALQOAPRPSG---RLQERAEQVSRALQPALDSYFHP 667
Db 428 RPGLARCLPVIELDDDDVLAAPARHLPPPTATGGTERLTDVRELRDTELRLTAEL---AP 484
Qy 668 GTPAPGRG-----VGPAGPGA 684
Db 485 AAPGTGRAALRRPLTAVGELARGGA 509

RESULT 15
O86440 PRELIMINARY; PRT; 361 AA.
AC O86440;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE XCPY PROTEIN.
GN XCPY.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID-303;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=WCS358;
RA de Groot A., Gerritse G., Lazdunski A., Filloux A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; X81085; CAA56986.1; -
SQ SEQUENCE 361 AA; 40565 MW; C364A65E526F998E CRC64;

Query Match 3.0%; Score 111; DB 2; Length 361;
Best Local Similarity 21.4%; Pred. No. 0.62; 113; Indels 194; Gaps 24;
Matches 92; Conservative 31; Mismatches 113; Indels 194; Gaps 24;

Qy 192 PALPWL-----NVSADGONVHLVNVSEEQHGLSLY 223
Db 10 PARFWLLLRPGSPDWLLVEGGIAQRQOGEPANLQA---RVALIVAEHCSHF----- 61
Qy 224 WNOVGPP---KPRWHKMLTGPQII---TLNHTDLVPCLCIQVWPLEPDSVRTNICPRE 277
Db 62 --QLPAPPGLRREW-----PLLEDRLQSADEVVCGCLG-----RE 97
Qy 278 DPAHQNLWQAARLRLTLQ-----SWLLDAPCS-----LPAE---AALCWR 316
Db 98 -----AGQVRLLTVARQHLEGNRGQCVKMGVSVECCWAEFQLLPAPPGVAVCMQ 147
Qy 317 APGDPCQPLVPP-----LSWENVTVDVNSSEKLOLO---ECLWADSLGPKDDVLL 366
Db 148 KPGFSLMKGLSEGAGERWLAWPDLGVDPVPEPQLRVPLEGRWPDELAPLDSLPGLF 207
Qy 367 ETRGPQDNRSICALEPSGCTSLPSKASTRAARLGEYLLQDLQSGCLOL---WDDDLGAL 423
Db 208 EQRPP---RTFVG-----PSRAQRQ-----LL-----AACLLLLACTW---GGL 239
Qy 424 WACPMCKYIHKRWALVWLACLLFRAALSILLKKDHAKGWLRLKQDVRSGAAAGRAA 483
Db 240 W-----LAQWRQVQL-----W-----QAQ 254
Qy 484 LLLYSADDSGFERLVGAL-----ASALCQLPLRVAVDLWSRELSAOGPVAV-----FH 532
Db 255 VLAVTGEQAGPRQAAQALKRLREAELOQQLRVRLADLQGRLOAWLQAHGWRLOAVRFD 314
Qy 533 AQRQTLOEQ 542
Db 315 GQRWHVRVEG 324

Search completed: May 19, 2002, 15:19:16
Job time: 6752 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:10:14 ; Search time 43.31 Seconds  
(without alignments)  
388.012 Million cell updates/sec

Title: US-09-608-918-10  
Perfect score: 3692  
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 688

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305.5	8.3	617	US-09-188-930-303	Sequence 303, Appl
2	176	4.8	866	US-08-620-694A-10	Sequence 10, Appl
3	176	4.8	866	US-09-022-255-10	Sequence 10, Appl
4	176	4.8	866	US-09-022-696-10	Sequence 10, Appl
5	176	4.8	866	US-08-978-773-4	Sequence 4, Appl
6	176	4.8	866	US-09-022-253-10	Sequence 10, Appl
7	176	4.8	866	US-09-022-260-10	Sequence 10, Appl
8	176	4.8	866	US-09-022-259-10	Sequence 10, Appl
9	176	4.8	866	US-09-022-257-10	Sequence 10, Appl
10	166	4.5	864	US-08-620-694A-2	Sequence 2, Appl
11	166	4.5	864	US-09-022-255-2	Sequence 2, Appl
12	166	4.5	864	US-09-022-696-2	Sequence 2, Appl
13	166	4.5	864	US-08-978-773-2	Sequence 2, Appl
14	166	4.5	864	US-09-022-253-2	Sequence 2, Appl
15	166	4.5	864	US-09-022-260-2	Sequence 2, Appl
16	166	4.5	864	US-09-022-259-2	Sequence 2, Appl
17	166	4.5	864	US-09-022-257-2	Sequence 2, Appl
18	147	4.0	330	US-09-188-930-125	Sequence 125, App
19	126.5	3.4	1711	US-08-342-930-2	Sequence 2, Appl
20	113	3.1	4928	US-09-036-987A-5	Sequence 5, Appl
21	113	3.1	4928	US-09-370-700-5	Sequence 5, Appl
22	109.5	3.0	428	US-09-423-340-4	Sequence 4, Appl
23	104	2.8	3724	US-08-804-227C-10	Sequence 10, Appl
24	104	2.8	3724	US-08-804-198-4	Sequence 4, Appl
25	103	2.8	6095	US-09-144-085-2	Sequence 2, Appl
26	102.5	2.8	1785	US-09-341-587-3	Sequence 3, Appl
27	102	2.8	1208	US-09-463-702A-2	Sequence 2, Appl

28	102	2.8	2629	2	US-08-751-189-4	Sequence 4, Appl
29	102	2.8	2629	2	US-09-060-836-4	Sequence 4, Appl
30	102	2.8	2629	4	US-09-184-445-4	Sequence 4, Appl
31	101.5	2.7	469	3	US-08-753-007A-8	Sequence 8, Appl
32	101.5	2.7	469	4	US-09-398-496-8	Sequence 8, Appl
33	100.5	2.7	405	2	US-08-483-151-4	Sequence 4, Appl
34	100	2.7	5588	4	US-09-036-987A-6	Sequence 6, Appl
35	100	2.7	5588	4	US-09-370-700-6	Sequence 6, Appl
36	98.5	2.7	647	3	US-08-753-007A-32	Sequence 32, Appl
37	98.5	2.7	647	4	US-09-398-496-32	Sequence 32, Appl
38	98.5	2.7	890	1	US-08-445-640-2	Sequence 2, Appl
39	98.5	2.7	890	3	US-08-170-558-2	Sequence 2, Appl
40	98.5	2.7	890	3	US-08-447-314-2	Sequence 2, Appl
41	98.5	2.7	890	3	US-08-445-461-2	Sequence 2, Appl
42	98.5	2.7	911	1	US-08-286-305A-1	Sequence 1, Appl
43	98.5	2.7	911	2	US-08-441-104A-1	Sequence 1, Appl
44	98.5	2.7	911	2	US-08-440-816A-1	Sequence 1, Appl
45	98.5	2.7	911	4	US-09-417-381A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-188-930-303  
; Sequence 303, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 303  
; LENGTH: 617  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-188-930-303

Query Match 8.3%; Score 305.5; DB 4; Length 617;  
Best Local Similarity 25.9%; Pred. No. 2.6e-20;  
Matches 130; Conservative 53; Mismatches 179; Indels 139; Gaps 18;

QY	237	KNLTGPOIITLNTDILVPCICIQVWPDPSPVTNICPFREDPRAH-QNLWQAAR-----	290
DB	183	KIVSGGHVVDLYEFLLPCMCIEASVYQEDTVRRKKCPQSWPEAFYGSDFWQSIRTDYS	242
QY	291	-----LRLLTQSLWLDAPCSLPFAEALCWRAPGDCQCPPLVPPLSWEN-----VTVDV	339
DB	243	QHONQVMALTLR-----CPLKLEASLCWRDPLTPCETLPNATAQEGEGWILENVDL	295
QY	340	N-----SSEKILQEQELWADSLGPKDDVLLLETRGPDQNRKSLCALPSPGCTSLPSKA	392
DB	296	HPQLCFKFSFENSSHVEC--PHQSGSLPSTVSMDTQAQQ-----LTLHFSSRTYATFSA	348
QY	393	STRAARLG-----EYLLQDLQSGQCLOLWDDDLGALW---ACP	427
DB	349	AWSDPGLGDPMPPPVYSISQTQGSVPVPTLDLIIFLROENCILVWRSDVHFAWKHVLCP	408
QY	428	MDKYIHKRWALVWLACLLFAAALSLLILKKDHAKGWLKLLKQDVPSGAARGAALLY	487
DB	409	DD-----APYPTQLLL-----RSLGSGRTRPVLLJH	434
QY	488	SADDSGFERLYGALASALCQLPLRVA-----VDLWSRRELSAQGPVAMFHAORROTL	539

```

Db 435 AADSEARRLVGALAE-----LRTALGGGRDVIIVDLWEGTHVARIGPLPWLWAARVA 489
QY 540 QEGGVVVVLFSPGAVALCSEWLQDGVGPGAHGPHDAFRASLSCVLPDFLOGRAPGSYVG 599
Db 490 REQGTVLLWN-----CAGPSTACSGDPOAASLRTL-----CAAPRPLLL 530
QY 600 ACFDRLLHPDAVPALRTVPVFWLPSPOLPFLGALQ-QPRAPRSG-----RLQERAE 650
Db 531 AYFSRLCAKGDIDPRPRALPRYLLRDLPRLLRALDAQPATLASSWSHLGAKRCLKNRLE 590
QY 651 QVS-RALQPALDSYFHPGTP 670
Db 591 QCHLLEAKDDYQGSTNSP 611

RESULT 2
US-08-620-694A-10
; Sequence 10, Application US/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION: 435
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-620-694A-10

```

Query Match 4.8%; Score 176; DB 2; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

```

QY 237 KNLITL---GPQIITLNLN---HT---DLVPLCLCIQVWPLEPDS-----VRTN-- 271
Db 66 RNLTPSPKDLQIQLHFAHTQQGDLFPVAHIE-WTLOTDAISILYGAELSVLQNTNER 124

```

```

QY 272 IC---PFREDPRAHQLWQAARLLTLTLOSLLDAPCS-----LPAEALCWRAPCGD 321
Db 125 LCVRELEFSLKRRHRRWR-----FTSFHVVDPDQYEVTVHLPKP-----IPDGD 172
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOLQOECLWADSLGPKL 360
Db 173 PNHQSNFLVDPCEHARMKVTPCMSSGSLWDPNITVETLEAQLRVSTFLWNES----- 227
QY 361 DVLLELTRGPQDNRSICALPESGCTS---LPSKASTRAARLGEY-----LLQDLQ 408
Db 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTLLRNK 274
QY 409 S-----GQCQLQWDDDL--GALWACP-----MDKYTHKRWALWLA 442
Db 275 GCCRHQVQIQPFSSCL---NDCLRHSATVSCPEMPDTPPEIPDYMLWVY--WFITGIS 329
QY 443 CLIFAAALSLLILLKKDHAKGWLRLKQDVR--SGAAA-----RGRAALLYSADDS 492
Db 330 ILLVGSVILLIVCMTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLKPRKVIISADHP 389
QY 493 GFERLVGALAS-ALCQLPLRVAVDLWSRRELSAQGPVAMFHAQRROTLOEGGVVVLIFSP 551
Db 390 LYVDVVVKFAQFLLTACGTEVALDLLEEQAISEAGVMTWVGRKQKQEWVESNKIIVLCR 449
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCVLPDFLOGRAPGSYVGACF 602
Db 450 GTRA---KW--QALLGRGAPVRLRCDHGKPGVGLFTAAAMNMLPDKRPACFGTVVVCYF 504
QY 603 DRLLHPDAVPALRTVPVFWLPSPOLPFLGALQOQPRAPSRGRLOQERAEQVSRALQPALDS 662
Db 505 SEVSCGDGVDPLFGAAPRYPLMDRFEVYFRIODLEMFOFGRH-HRVGELSG-----DN 557
QY 663 YEHPPG 668
Db 558 YLRSPG 563

RESULT 3
US-09-022-255-10
; Sequence 10, Application US/09022255
; Patent No. 6072033
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,255
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/620,694
; FILING DATE: 21 MARCH 1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne

```

REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-255-10

Query Match 4.8%; Score 176; DB 3; Length 866;  
Best Local Similarity 21.8%; Pred. No. 8.3e-08;  
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNTL--GPOIITLN---HT---DLVPCLCIOVWPLEPDS-----VRTN-- 271  
DB 66 RNLTPSPFKDLOIQHFAHTQOGLFPVAHIE-WLQTDASILYLEGAELSVLQNTNER 124  
QY 272 IC---PFREDPRAHONLQWAAARLLTLQSWLLDAPCS-----LPAAALCWRAPGGD 321  
DB 125 LCVREFEFLSKLRHHRRWR-----FTFSHEVVDPDQYEYVTHLKP-----IPDGD 172  
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOEQECLWADSLGPKL 360  
DB 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227  
QY 361 DDVLLLETRGPQDNRSCLALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408  
DB 228 THYQILLTSPFH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRNK 274  
QY 409 S-----GQCQLWDDDL--GALWACP-----MDKYTHKRWALVWLA 442  
DB 275 GCRHVOIQPFSSCL---NDCLRHSAIVSCPEMDTPPEIPDYMLVWY--WFTGIS 329  
QY 443 CLLFAAALSILLKKDHAKGWLRLKQDVR--SGAAA-----RGAALLYSADDS 492  
DB 330 ILLVGSVILLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPKRWIIYSADHP 389  
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAOGPVANFHAQRRTLOEGGVVLLFSP 551  
DB 390 LYVDVWLKFAQLLACGTEVALDLEEQAISEAGVMTWVGROKQEMVESNSKIIVLCRSR 449  
QY 552 GAVALCSEWLQDVGSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSGYVACF 602  
DB 450 GTRA---KW--QALLGRGAPVRLKCDHGKPVGDLFTAAMNMLLPDKRPACFTYVVCYF 504  
QY 603 DRLLHPDAPVLPALFRTVPVFTLPSQLPDFLQALQOPRPRSGRLQERAEQVSRALQPALDS 662  
DB 505 SEVSCDGDVDPDLFGAAPRYLMDRFEVYFRIQDLEMFQGRM-HRVGELSG-----DN 557  
QY 663 YHPPG 668  
DB 558 YLRSPG 563

RESULT 4  
US-09-022-696-10  
Sequence 10, Application US/09022696  
Patent No. 6072037  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle

STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-696-10

Query Match 4.8%; Score 176; DB 3; Length 866;  
Best Local Similarity 21.8%; Pred. No. 8.3e-08;  
Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNTL--GPOIITLN---HT---DLVPCLCIOVWPLEPDS-----VRTN-- 271  
DB 66 RNLTPSPFKDLOIQHFAHTQOGLFPVAHIE-WLQTDASILYLEGAELSVLQNTNER 124  
QY 272 IC---PFREDPRAHONLQWAAARLLTLQSWLLDAPCS-----LPAAALCWRAPGGD 321  
DB 125 LCVREFEFLSKLRHHRRWR-----FTFSHEVVDPDQYEYVTHLKP-----IPDGD 172  
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKLOEQECLWADSLGPKL 360  
DB 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227  
QY 361 DDVLLLETRGPQDNRSCLALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408  
DB 228 THYQILLTSPFH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRNK 274  
QY 409 S-----GQCQLWDDDL--GALWACP-----MDKYTHKRWALVWLA 442  
DB 275 GCRHVOIQPFSSCL---NDCLRHSAIVSCPEMDTPPEIPDYMLVWY--WFTGIS 329  
QY 443 CLLFAAALSILLKKDHAKGWLRLKQDVR--SGAAA-----RGAALLYSADDS 492  
DB 330 ILLVGSVILLIVCMTWRLAGPSEKYSDDTKYTDGLPAADLIPPLKPKRWIIYSADHP 389  
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAOGPVANFHAQRRTLOEGGVVLLFSP 551  
DB 390 LYVDVWLKFAQLLACGTEVALDLEEQAISEAGVMTWVGROKQEMVESNSKIIVLCRSR 449  
QY 552 GAVALCSEWLQDVGSGPGA-----HGP--HDAFRASLSCVLPDFLQGRAPGSGYVACF 602  
DB 450 GTRA---KW--QALLGRGAPVRLKCDHGKPVGDLFTAAMNMLLPDKRPACFTYVVCYF 504  
QY 603 DRLLHPDAPVLPALFRTVPVFTLPSQLPDFLQALQOPRPRSGRLQERAEQVSRALQPALDS 662



```

Db 275 GCCRHQVQIQFTSSCL---NDCLRHSATVSCPEMDTPEPIPDYMLWY--WFTIGIS 329
QY 443 CLLFAAALSILLLKKDHAKGWLRLLLKQDVR--SGAAA-----RGRAALLLYSADDS 492
Db 330 ILLVGSVILLIVCTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLKPKRWIIYSADHP 389
QY 493 GFERLVGALAS-ALCQLPLRVAVDILMSRRRELSAOGPVAMFHAAORRTOLEGGVVVLLFSP 551
Db 390 LYDVLKFAQLLLTACGTEVALDLEEQAISEAGVMTWVGKGQEMVESKIIIVLCSR 449
QY 552 GAVALCSENLQDVGSPGA-----HGP--HDAFRASLSCLVLPDFLQGRAGSGSYVGACF 602
Db 450 GTRA---KW--QALLRGAPVRLRCDHGKPVGDLFTAAAMNMLLPDFKRPACGTYVVCYF 504
QY 603 DRLLHPDAVPALEFRTVPVFTPLSQLPDFLGAQQAPRPSGRQLQERAEQVSRALOPALDS 662
Db 505 SEVSCDGDVDFLGAAPRYPLMDRFEENVYRIODLEMFQGRM-HRVGELSG-----DN 557
QY 663 YFHPPG 668
Db 558 YLRSPG 563

RESULT 6
US-09-022-253-10
; Sequence 10, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, William
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694
; FILING DATE: 21-MARCH-1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

US-09-022-253-10

Query Match 4.8%; Score 176; DB 3; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOIITLN---HT---DLVPCICIQVWPLEPDS-----VRN-- 271  
DB 66 RNLTPSSPKDLQQLHFAHTQOQGLFPVAHIE-WTQTDASILYEGAELSVLQNTN 124  
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAEALCWRAPGD 321  
DB 125 LCVREFEELSKLRHHRRWR-----FTFSHFVDPDQOEYEVTVHHLKP 172  
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKIQCEIWLADSLGPK 360  
DB 173 PNHQSKNFLVPDCEHARMKVTTPCMSSGSLWDPNITVETLEAHLRVSTLWNES----- 227  
QY 361 DDVLLLETRGPDNRSLCALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408  
DB 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRLNK 274  
QY 409 S-----GQCLQLWDDDL--GALWACP-----MDKYIHKRWALVWLA 442  
DB 275 GCRHVOVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPEIPDYMPLWVY--WFTGIS 329  
QY 443 CLLFPAALSLILLKKDHAKGWLRLLLKQDVR--SGAAA-----RGRAALLYSADDS 492  
DB 330 ILLVGSVILLIVCMWRLAGPGSEKYSDDTKYTDGLPAADLIPPLKPKRWIIYSADHP 389  
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFWFAHQRROTLOEGGVVLLFSP 551  
DB 390 LYVDVVLKFAQFLTLTACGTEVALDLLEQAISEAGVMTWVGROKQEMVESNSKIIVLCR 449  
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCLVLPDFLOGRAPGSVYVACF 602  
DB 450 GTRA---KW--QALLGRGAPVRLRCDHGKPVGDLEFTAAMNMILPDFKRPACFGYVVCYF 504  
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAPRSGRQERAEQVSRALQPALDS 662  
DB 505 SEVSCDGDVDFLEGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG-----DN 557  
QY 663 YFHPGP 668  
DB 558 YLRSPG 563

RESULT 7

US-09-022-260-10

Sequence 10, Application US/09022260

Patent No. 6100235

GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie

APPLICANT: Fanslow, William

TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh

OPERATING SYSTEM: Apple Operating System 7.5.5

SOFTWARE: Microsoft Word for Apple, Version 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/022,260

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION NUMBER: 08/620,694

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/410,535

FILING DATE: 23 MARCH 1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,695

REFERENCE/DOCKET NUMBER: 2617-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)587-0430

TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 866 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-022-260-10

Query Match 4.8%; Score 176; DB 3; Length 866;

Best Local Similarity 21.8%; Pred. No. 8.3e-08;

Matches 119; Conservative 76; Mismatches 189; Indels 162; Gaps 32;

QY 237 KNL--GPOIITLN---HT---DLVPCICIQVWPLEPDS-----VRN-- 271  
DB 66 RNLTPSSPKDLQQLHFAHTQOQGLFPVAHIE-WTQTDASILYEGAELSVLQNTN 124  
QY 272 IC---PFREDPRAHONLWQAARLLTLQSWLLDAPCS-----LPAEALCWRAPGD 321  
DB 125 LCVREFEELSKLRHHRRWR-----FTFSHFVDPDQOEYEVTVHHLKP 172  
QY 322 PCQP-----LVPPLS-----WE-NVTVDVNSSEKIQCEIWLADSLGPK 360  
DB 173 PNHQSKNFLVPDCEHARMKVTTPCMSSGSLWDPNITVETLEAHLRVSTLWNES----- 227  
QY 361 DDVLLLETRGPDNRSLCALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408  
DB 228 THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEEFHORSNVTTLRLNK 274  
QY 409 S-----GQCLQLWDDDL--GALWACP-----MDKYIHKRWALVWLA 442  
DB 275 GCRHVOVQIOFFSSCL---NDCLRHSATVSCPEMPDTPPEIPDYMPLWVY--WFTGIS 329  
QY 443 CLLFPAALSLILLKKDHAKGWLRLLLKQDVR--SGAAA-----RGRAALLYSADDS 492  
DB 330 ILLVGSVILLIVCMWRLAGPGSEKYSDDTKYTDGLPAADLIPPLKPKRWIIYSADHP 389  
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFWFAHQRROTLOEGGVVLLFSP 551  
DB 390 LYVDVVLKFAQFLTLTACGTEVALDLLEQAISEAGVMTWVGROKQEMVESNSKIIVLCR 449  
QY 552 GAVALCSEWLQDGVSGPGA-----HGP--HDAFRASLSCLVLPDFLOGRAPGSVYVACF 602  
DB 450 GTRA---KW--QALLGRGAPVRLRCDHGKPVGDLEFTAAMNMILPDFKRPACFGYVVCYF 504  
QY 603 DRLLHPDAVPALFRTVPVFTLPSQLPDFLGAQQPRAPRSGRQERAEQVSRALQPALDS 662  
DB 505 SEVSCDGDVDFLEGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG-----DN 557  
QY 663 YFHPGP 668  
DB 558 YLRSPG 563

RESULT 8

US-09-022-259-10

Sequence 10, Application US/09022259



QY 237 KNT--GPOIITL---HT---DLVPCLCIQVWPLEPDS-----VRN-- 271  
Db 66 RNLTPSSPKDLQLHFAHTQOGDLFPVAHIE-WTLOTDSASILYLEGAELSVQLNTNR 124  
QY 272 IC---PFREDPRAHQLWQAARLLTLQSWLLDAPCS-----LPAEALCWAPAGD 321  
Db 125 LCVREFELSKLRHHRRWR-----FTFSHFVYDPOEYEVTVHHLKP-----IPGD 172  
QY 322 PCOP-----LVPPLS-----WE-NVTVDVNSSEKLOEQCLWADSLGPKL 360  
Db 173 PNHQSKNFLVPDCEHARMKVTPCMSSGSLWDPNITVETLEAHQLRVSTLWNES----- 227  
QY 361 DDVLLLETRGPQDNRSICALEPSGCTS-----LPSKASTRAARLGEY-----LLQDLQ 408  
Db 228 THYQILLTSPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTLTLENK 274  
QY 409 S-----GOCQLQWDDDL---GALWACP-----MDKYTHKRWALWLA 442  
Db 275 GCRHQVQLOPFSSCL---NCLRHSAIVSCPEMDTPPIPDYMLVY--WFTGIS 329  
QY 443 CILFAAALSILLLKKDHAKGWLRLKQDVR---SGAAA-----RGAALLLSYSDS 492  
Db 330 ILLVGSVILLIIVCMTRWLAGPSEKYSDDTKYTDGLPAADLIPPLPKPRKVIISADHP 389  
QY 493 GFERLVGALAS-ALCOLPLRVAVDLWSRRELSAQGPVAFHQAQRQTLOEGGVVLLFSP 551  
Db 390 LXDVVYKFAQPLLACGTAGTEVALDLLEQAISEAGVMTWVGROKQEVESNKIIVLCR 449  
QY 552 GAVALCSEWQGVSGPA-----HGP---HDAFRASISCVLPDFLQGRAPSYGACF 602  
Db 450 GTRA---KW-QALLGRGAPVRLRCDHGXPGVGLDFTAAMNMLIPDFKRPACFTGVVYCF 504  
QY 603 DRLLHPDPAVLPFRVTVLPSQLPDFLGAQQPRAPRSGRLOERAEQVSRALQPADS 662  
Db 505 SEVSCGDVDPDLFGAAPRYPLMDRFEFVYFRIQDLEMFQGRM-HRVGELSG-----DN 557  
QY 663 YFHPG 668  
Db 558 YLRSPG 563

RESULT 10

US-08-620-694A-2  
; Sequence 2, Application US/08620694A  
; Patent No. 5869286

GENERAL INFORMATION:

; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,694A  
; FILING DATE: 21 MARCH 1996  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/538,765  
; FILING DATE: 7 AUGUST 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/410,535

; FILING DATE: 23 MARCH 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 864 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-620-694A-2

Query Match 4.58; Score 166; DB 2; Length 864;  
Best Local Similarity 20.4%; Pred. No. 7.4e-07;  
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW---LNVASADGNVHLVNL---VSEOHFGSLYWNVOGPP--KPRW--HKNL 239  
Db 14 PALGWLILLNLVAPGRASPRLLDFPAPVCAQE--GLSC---KVNSTCLDDSIHKNL 68  
QY 240 T--GPQIITL-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273  
Db 69 TPSPKNIYINLSVSTQHGELVPVLHVE-WTLOTDSASILYLEGAELSVQLNTNERLCV 127  
QY 274 --PREDPRAHONLWQAARLLTLQSWLLDAPCSLPAEA-----ALCWAPAGDP--- 322  
Db 128 KFQFLSMLQHRKRWR-----FSFSHFVVD---PQOEYEVTVHHLPKPIPDGPNHK 176  
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLOEQCLWADSLGPKLDDVL 364  
Db 177 SKIIFVPCDESKMKMTTSCVSSGSLWDPNITVETLQHLRVFTLWNEST-PYQ---V 232  
QY 365 LLETGPDQNRSLCALEPSPGCTSL-----PSKASTRAARLGEYLLQDLSGOCQLQWDD 418  
Db 233 LLESFSDSENHS-----CFDVVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274  
QY 419 DLGALWAC-----PMDKYIHKRW-----LV 439  
Db 275 ---WCCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PLWVYGLIT 328  
QY 440 WLACLLFAAALSILLLKKDHAKGWLRLKQDVRSGA-----SGDOEKHGGDSKINGILPVADTTPP 375  
Db 329 LIALLVGSVIVLICTMT-----W-RL-----AVAL-CSEWLODGVSGPAGHPDAFRAS 580  
QY 477 AARGAALLLYSADDSGFERLVGALASAL-COLPLRVAVDLWSRRELSAQGPVAFHQAQR 535  
Db 376 PLRPRKVIWIVYSADHPLYVEVVLKFAQFLITACGTAEVALDLLEEQVISEGVMTWVSQK 435  
QY 536 RQTLOEGGVVLLFSPG-----AVAL-CSEWLODGVSGPAGHPDAFRAS 580  
Db 436 QEMVESNKIILCSRGTOAKWKAILGWAEPVQLRCDHMKPAG-----DLFTAA 485  
QY 581 LSCVLPDFLOGRAPGSYVACFDRLLHPDPAVLPFRVTVVFTLL/SOLPDFLGAQQPRAP 640  
Db 486 MNMLPDKRPACGFTGVVYFSGICSERDVPDLNITSRYPLMDRFEVYFRIQDLEMF 545  
QY 641 RSGRLQERAEQVSRALQPADLSYFHP 667  
Db 546 EPGRMHHVRELGTG-----DNYLQSP 565

RESULT 11

US-09-022-255-2  
; Sequence 2, Application US/09022255  
; Patent No. 6072033  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin

APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/620,694  
FILING DATE: 21 MARCH 1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-255-2

Query Match 4.5%; Score 166; DB 3; Length 864;  
Best Local Similarity 20.4%; Pred. No. 7.4e-07;  
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

QY 192 PALPW-----LNVSGDGNVHLVILN-----VSEQHFGSLYWNQVQGP--KPRW--HKNL 239  
Db 14 PALGWLILLNLVLAAGRSPRLDFFAPVCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68  
QY 240 T--GPOIITN-----HTDLVPLCIQWPLEPDS-----VRTN--IC- 273  
Db 69 TPSSPKNIYINLSVSTQHGELVPVHLVE-WTLQTDASILYLEGAELSVLQNLNTERLCV 127  
QY 274 --PFREDPRAHONLQAARLLTLQSLWLDAPCSLPAAE-----ALCNRAPGCDP--- 322  
Db 128 KQQLSLMLQHRRWR-----FSFHFVVD-----PGQYEYTVVHLHPKPIPGDGNHK 176  
QY 323 -----COPLVPPPLS-----WE-NVTVDVNSSEKLQLOECLWADSLGPKLDDVL 364  
Db 177 SKIIFVPDCEDSKMKMTTSCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232  
QY 365 LLETGCPQDNRSICALPEGCTSL-----PSKASTFAARLGEYLLQDLQSGCQLQWD 418  
Db 233 LLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274  
QY 419 DLGALWAC-----PMDKYITHKWA---LV 439  
Db 275 -----WCCHHHVQVQFFFSCLNDCLRHAVTVPCVISNTVTPKPVADYI-PLWVYGLIT 328  
QY 440 WLACLLFAAALSILLKKDKHAKGWLRLKQDVRSGA----- 476

Db 329 LIALLVGVSIVILICMT-----W-RL-----SGAQEKHGDDSKINGILPVADLTPP 375  
QY 477 AARGAALLLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSSAQGPVAMFHAQR 535  
Db 376 PLRPRKWIWIVSADHPLVVEVVLKFAQLITACGTEVALDLLEQVISEVGVMTWVSQK 435  
QY 536 RQTLQEGGVVLLFSPG-----AVAL-CSENLQDGVSGPGAHGPHDAFRAS 580  
Db 436 QEMVESNSKIILCSRGTOAKWAILGWAEPVQLRCDHWKPAG-----DLFTAA 485  
QY 581 LSCVLPDFLOGRAPGSGVYACFDFRLLHDDAVPALFRTVPVFTLPSQLPDFLGCALQOPRAP 640  
Db 486 MNMILPDFKRPACFGTYVYVCFSGICSRDPDPLFNITSRYPLMDRFEVVERIODLEMF 545  
QY 641 RSGRLQERAEQVSRALQPALDSYFHP 667  
Db 546 EPGRMHVRELTG-----DNYLQSP 565

RESULT 12  
US-09-022-696-2  
Sequence 2, Application US/09022696  
Patent No. 6072037  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-696-2

Query Match 4.5%; Score 166; DB 3; Length 864;  
Best Local Similarity 20.4%; Pred. No. 7.4e-07;  
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

```
QY 192 PALPW-----LNVSADGDNVHLVNLN-----VSEQHFGSLYWNQVQPP--KPRW--HKNL 239
Db 14 PALGWLILLNLVAPGRASPRLLDFFAPVCAQE--GLSC---RVKNSTCLDDSWIHPKKNL 68
QY 240 T--GPOIITLN-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273
Db 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLOTDASILYLEGAELSVLQNTNERLCV 127
QY 274 ---PFREDPRAHONLWOAARLLTLQSWLLDAPCSLPAAE-----ALCWRAPEGDP--- 322
Db 128 KFQFLSMQLQHRKRW-----FSFSHFVVD-----PGQYEVTVVHHLPKPIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLQLECLWADSLGPKLDDVL 364
Db 177 SKIIFVPCDESKMKMTTSCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGQNRSLCALEPGCGTSL-----PSKASTRAARLGEYLQDLQSGQCLQLWDD 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----WCCHHHVQVQPFSSCLNCLRHAVTPCPVISNTTVPKPVADYI-PLWVYGLIT 328
Db 275 -----WCLACLLFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
QY 440 WLACLLFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
Db 329 LIAILLVGSVIVLIICMT-----W-RL-----SGAQEKHGDDSKINGILPVADITPP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSSAQGPVAFHQAOR 535
Db 376 PLRPRKVMIVYSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWVSROK 435
QY 536 RQTLQEGGVVLLFSPG-----AVAL-CSEWLODGVSGGAGHPHDAFRAS 580
Db 436 QEMVESNSKIIILCSRGTOAKWKAILGWAEPVQLRCDHMKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAGSYGACFDRLLHPDAVPALFRTVPVFTPLPSQLPDFLGAQQPRAP 640
Db 486 MNMILPDFKRPACFGTYVYVYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIOQLEMF 545
QY 641 RSGRLQERAEQVSRLALQALPDSYFHP 667
Db 546 EPRGHHVRELGT-----DNYLQSP 565
```

## RESULT 13

```
US-08-978-773-2
; Sequence 2, Application US/08978773
; Patent No. 6083906
; GENERAL INFORMATION:
; APPLICANT: Troutt, Anthony
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple PowerMacintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,773
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/052,525
; FILING DATE: 27 NOVEMBER 1996
```

```
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2623-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-978-773-2
```

Query Match 4.5%; Score 166; DB 3; Length 864;

Best Local Similarity 20.4%; Pred. No. 7.4e-07; Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;

```
QY 192 PALPW-----LNVSADGDNVHLVNLN-----VSEQHFGSLYWNQVQPP--KPRW--HKNL 239
Db 14 PALGWLILLNLVAPGRASPRLLDFFAPVCAQE--GLSC---RVKNSTCLDDSWIHPKKNL 68
QY 240 T--GPOIITLN-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273
Db 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLOTDASILYLEGAELSVLQNTNERLCV 127
QY 274 ---PFREDPRAHONLWOAARLLTLQSWLLDAPCSLPAAE-----ALCWRAPEGDP--- 322
Db 128 KFQFLSMQLQHRKRW-----FSFSHFVVD-----PGQYEVTVVHHLPKPIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLQLECLWADSLGPKLDDVL 364
Db 177 SKIIFVPCDESKMKMTTSCVSSGSLWDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGQNRSLCALEPGCGTSL-----PSKASTRAARLGEYLQDLQSGQCLQLWDD 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----WCCHHHVQVQPFSSCLNCLRHAVTPCPVISNTTVPKPVADYI-PLWVYGLIT 328
Db 275 -----WCLACLLFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
QY 440 WLACLLFAAALSILLKLDHAKGWLRLKQDVRSGA----- 476
Db 329 LIAILLVGSVIVLIICMT-----W-RL-----SGAQEKHGDDSKINGILPVADITPP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSSAQGPVAFHQAOR 535
Db 376 PLRPRKVMIVYSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTWVSROK 435
QY 536 RQTLQEGGVVLLFSPG-----AVAL-CSEWLODGVSGGAGHPHDAFRAS 580
Db 436 QEMVESNSKIIILCSRGTOAKWKAILGWAEPVQLRCDHMKPAG-----DLFTAA 485
QY 581 LSCVLPDFLOGRAGSYGACFDRLLHPDAVPALFRTVPVFTPLPSQLPDFLGAQQPRAP 640
Db 486 MNMILPDFKRPACFGTYVYVYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIOQLEMF 545
QY 641 RSGRLQERAEQVSRLALQALPDSYFHP 667
Db 546 EPRGHHVRELGT-----DNYLQSP 565
```

## RESULT 14

```
US-09-022-253-2
; Sequence 2, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
```



```
Matches 128; Conservative 77; Mismatches 196; Indels 226; Gaps 35;
QY 192 PALPW-----LNYSDAGDNVHLVNLN-----VSEEQHFGLSLYWNQVOGPP--KPRW--HKNL 239
Db 14 PALGWLLLLLNLVLA GRASPRLLDFPAPVCAQE--GLSC---RVKNSTCLDDSDWIHPKNL 68
QY 240 T--GPOIITLN-----HTDLVPCICIQWPLEPDS-----V--VRTN--IC- 273
Db 69 TPSSPKNIYINLSVSTQHGELVPVLHVE-WTLOTDASILYLEGAELSVLQLTNTERLCV 127
QY 274 --PFREDPRAHQNLQAARLLTLTQSLWLDAPCSLPAAE-----ALCWRAPGGDP--- 322
Db 128 KFOFLSMLQHRRKRWR-----FSFSHFVVD-----PGQEVETVHHLPKPIPDGDPNHK 176
QY 323 -----COPLVPPLS-----WE-NVTVDVNSSEKLQLOECLWADSLGPKDDVL 364
Db 177 SKIIFVPDCEDSKMKMTTSCVSSGSLMDPNITVETLDTQHLRVDFTLWNEST-PYQ---V 232
QY 365 LLETGPGQDNRSICALEPSCGCTSL-----PSKASTRAARLGEVLLQDLQSGQCQLQWDD 418
Db 233 LLESFSDSENHS-----CFDVVKQIFAPROEEFHQRANVTFTLSKFH----- 274
QY 419 DLGALWAC-----WCCHHHVQVQPFSSCLNDCLRHAHTVPCPVISNTTVPKPVADYI-PLWVYGLIT 328
Db 275 -----WCCHHHVQVQPFSSCLNDCLRHAHTVPCPVISNTTVPKPVADYI-PLWVYGLIT 328
QY 440 WLACLLFAAALSILLKKDKHAKGWLRLKQDVRSGA----- 476
Db 329 LTAILLVGSVIVLIICMT-----W-RL-----SGADQEKHGDDSKINGILPVADLTTP 375
QY 477 AARGRAALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAQGPVAVFHAQR 535
Db 376 PLRPRKVIWIVYSADHPLYVEWLKFAQLITACGTEVALDLEEQVISEVGMVTWVSQRK 435
QY 536 ROTLORGGVVVLLGSPG-----AVAL-CSEWLQDGVSGPGAGHPHDAFRAS 580
Db 436 QEWVESNSKIIILCSRGTOAKWKAILGWAEPVQLRCDHWKPG-----DLFTAA 485
QY 581 LSCVLPDFLQGRAPGSYGACFDRLLHPDAPALFTVPVFTLPFSQLPDFLGALQOPRAP 640
Db 486 MNMILPDFKRPACFGTYVVCYFGICSERDVPDLFNITSRYPLMDRFEVYFRIQDLEMF 545
QY 641 RSGRLQERAEQVSRALQPALDSTFHPP 667
Db 546 EPGRMHHVRELGT-----DNYLQSP 565
```

Search completed: May 19, 2002, 14:10:18  
Job time: 6375 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 12:24:52 ; Search time 63.07 seconds  
(without alignments)  
1054.286 Million cell updates/sec

Title: US-09-608-918-2

Perfect score: 3716

Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	137.5	3.7	3707	2	S18252
2	130.5	3.5	1366	2	S35985
3	115.5	3.1	745	2	H85840
4	114.5	3.1	745	2	C99995
5	114.5	3.1	1193	2	F83264
6	113.5	3.1	761	2	S20458
7	111.5	3.0	478	2	D75564
8	108	2.9	2629	2	T30987
9	107	2.9	719	2	T35189
10	107	2.9	825	1	A60386
11	107	2.9	1118	1	A49724
12	104.5	2.8	238	2	C70841
13	104.5	2.8	975	2	S33121
14	104.5	2.8	1711	1	A55148
15	104	2.8	1487	2	T02850
16	103.5	2.8	676	1	WMBEX6
17	103.5	2.8	828	2	AD0412
18	103.5	2.8	2055	2	T00093
19	103	2.8	4845	2	T31067
20	101.5	2.7	1151	2	S48431
21	101	2.7	274	2	A83583
22	100.5	2.7	709	2	F75584
23	100.5	2.7	741	2	T31164
24	100.5	2.7	753	1	WZBPE8
25	99	2.7	429	2	T36088
26	99	2.7	636	2	B83513
27	99	2.7	660	2	E98169
28	99	2.7	660	2	AB3118
29	99	2.7	679	2	B75262

30	98.5	2.7	724	2	T47149
31	98.5	2.7	741	2	B49555
32	98.5	2.7	1737	2	T00209
33	98	2.6	427	2	E87669
34	98	2.6	1487	1	EDBEF1
35	98	2.6	1487	1	EDBEF6
36	98	2.6	2205	1	MNVVRN
37	97.5	2.6	418	2	JC7588
38	97.5	2.6	617	2	F75484
39	97.5	2.6	698	2	A82593
40	97.5	2.6	850	2	JC5700
41	97.5	2.6	3739	2	T17410
42	97.5	2.6	4391	2	A38096
43	97	2.6	415	2	T46462
44	97	2.6	779	2	AG1978
45	97	2.6	2594	2	A35774

ALIGNMENTS

RESULT 1

S18252

heparan sulfate proteoglycan - mouse

N:Alternate names: perlecan

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999

C:Accession: S18252; A31917; B31917; S66460

R:Noonan, D.M.; Fulle, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.;

J. Biol. Chem. 266, 22939-22947, 1991

A:Title: The complete sequence of perlecan, a basement membrane heparan sulfate proteoglycan molecule.

A:Reference number: S18252; MUID:92078153

A:Accession: S18252

A:Molecule type: mRNA

A:Residues: 1-3707 <NOO>

A:Cross-references: EMBL:M77174; NID:920295; PIDN:AAA39911.1; PID:9200296

R:Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.;

J. Biol. Chem. 263, 16379-16387, 1988

A:Title: Identification of cDNA clones encoding different domains of the basement membrane heparan sulfate proteoglycan molecule.

A:Reference number: A92680; MUID:89034110

A:Accession: A31917

A:Molecule type: mRNA

A:Residues: 940-1601 <NO>

A:Cross-references: GB:J04054; NID:9200252; PIDN:AAA39899.1; PID:9200253

A:Accession: B31917

A:Molecule type: mRNA

A:Residues: 1870-2600 <NO3>

A:Cross-references: GB:J04055; NID:9200300; PIDN:AAA39912.1; PID:9200301

R:Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpl, R.

Eur. J. Biochem. 231, 551-556, 1995

A:Title: Structural properties of recombinant domain III-3 of perlecan containing a glycosaminoglycan chain.

A:Reference number: S66460; MUID:95377282

A:Accession: S66460

A:Molecule type: protein

A:Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>

C:Superfamily: LDL receptor ligand-binding repeat homology: EGF homology; laminin G repeat homology

C:Keywords: glycoprotein

F:195-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F:325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F:368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F:764-811/Domain: laminin-type EGF-like homology <LEG>

F:1159-1206/Domain: laminin-type EGF-like homology <LEG7>

F:1563-1610/Domain: laminin-type EGF-like homology <EGF>

F:1613-1668/Domain: laminin-type EGF-like homology <LEG8>

F:3163-3198/Domain: EGF homology <EGF>

F:3270-3423/Domain: laminin G repeat homology <EGF>

F:3464-3492/Domain: EGF homology <EGF7>

F:1256,1891,2336,2394,2427/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.7%; Score 137.5; DB 2; Length 3707;

Best Local Similarity 19.58; Pred. No. 0.13; Mismatches 235; Indels 365; Gaps 42;  
Matches 163; Conservative 73;

```
QY 4 PWFLLSLALGRSPVLSLERLVGPDAT-----HCSP-----GLSC----- 39
Db 272 PQFLPSPVPGS-----ACGPQASCHSGHCIPRDYLCDDGQEDCRDGSDELGCASPP 323
QY 40 -----RLWSDILCLPGDI-----VPAPGVLPATHTLQTEL 71
Db 324 PCBPNEFACENGHCALKLRCD-----GDFCEDRTDEANGSVKQPGVEGCTHFCQVST 378
QY 72 LR-----COKETDCDLCLRVAVHLAVHGHWEPEDEEKGGAADSGVEPRNASLOAQ 124
Db 379 NRCIPASFHDESDC-----PDRSDFEGCMPPQVVTTPQO-SIOA- 418
QY 125 VVLSFOAYPTARCVLLEVQVPAALVQFGSVGVVYDCFEAALGSEVRINWYTPRYEKE 184
Db 419 -----SRGQ-----TVTFCTVATGVPTPIINW-----R 441
QY 185 LNHTQOLPALPMLNVSADGNVHLVNLVSEOHFG-----LSLYW 224
Db 442 LN-WGHIPAPRPVMTSEGGRTLIIRDVKEADQAVTCEAMNSRGMVFGIPDGVLELV- 499
QY 225 NOVQGPCKPRWHKNLTGQIITLNHTDVLPCICIQVWPLEPDSVRTNCP-----FREDPR 280
Db 500 --PQPGCPDGHFYLE-----DSASCLPCFCFV-----TNVCOSSLRFRDQIR 541
QY 281 A---HONLWQAARLRLTLQSWLLDAPCSLPAEALCWAPGGPCOPPLVPPLSWENVTV 337
Db 542 LSPDQNDKFGVNTM-----PSQGPVPLSLSTQLOI 573
QY 338 DKVL-BFPLKGNPCVQVNSSEKLOLQCLWA-----DSLQ----- 374
Db 574 DPALQEFOL-----VDLSRRFLVHDFAFWALPKQFLGNKVDSYGGFLRYKRYELA 623
QY 375 -----PL-RDDVLL-----LETGRQDNRLCALPESGCTSLPSKASTRAARIGEYLLQ 422
Db 624 RGMLEPVQKPDVILVAGYRLHSGRGTPTHT-----PGTLNQRQVQLSE----- 666
QY 423 DLQSGOCLQLDWDLDLQALMACPMKYIHKRWALVWLACLLFA-AALSILILLKKDHAKAA 481
Db 667 -----EHWVHESGRVPQRAEMLOALASLEAVLLQTVYNTKMA 703
QY 482 AGRRAALLYSDSGFERLVGALASAL--COLPLRVAVDLWSRRELSAOGPVAMFHAQR 539
Db 704 SVGLSDIVM---DITVTHTTIHGRAHSVEBCRPIGVS-----GLSCESCDAHF----- 749
QY 540 RQTLQEGGVVLLFS-----PGAVALC-----SEWLDQGVSGPGAGHGPDAFR 582
Db 750 --TRVPGGYLGTCSGCNCGHASSCDPVYGHCLNCOHNTGEPQCDKCKPGFFG--DATK 805
QY 583 ASLSCVLPDLQGRAPGSYVGA-----CFDRLHHPDAVPALFRTPVPT-----LPSQLP 632
Db 806 ATATACRP-----CPCPIDASRRSDFTCF---LQDQATCDACAPGYTGRRCESCAP 856
QY 633 DFILGALQOP-----RAPRSGRLOERAEQ-----VSRLQPALDYSVFH 669
Db 857 GYEGNPIQPGKCRPTQETIVRCDERSGLTSGTCRCKNNVGRCLNCECSDGSF 912
```

## RESULT 2

T35985 probable large pro/Ala/Gly-rich protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T35985

R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1999

A:Reference number: Z21589

A:Accession: T35985

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1366 <MUR>

A:Cross-references: EMBL:AL035559; PIDN:CAB37473.1; GSPDB:GN00070; SCOEDB:SC9F2.06c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC9F2.06c

Query Match 3.5%; Score 130.5; DB 2: Length 1366;  
Best Local Similarity 22.4%; Pred. No. 0.13;  
Matches 175; Conservative 57; Mismatches 238; Indels 311; Gaps 42;

```
QY 21 LERLVGPQDATHCSGLSRLWSDILCLPGDIVAPGVLPATHTLQTELVRCKOKETOC 80
Db 675 LDGALADPDAGH---PLTIRLSEVRAALPGP--PAPVEVTRDAVETAYLDL----- 721
QY 81 DCLCLRVAVHLAVHGHWEPEDEEKGGA-----DSGVEEPRNASI-QAQVVLSSFOA 131
Db 722 -MCLRVATRLA-----DENGURGTAVRRLAKVSGVQVHEAARRSLPGCGGLDRS 771
QY 132 YPT-----ARCVLLE-VQVPAAL-----VQFGQSVG 156
Db 772 FETLPCGPAPARLGGGTGWAPVLAEGLFVPTGSGYRFAHEELADWIOGTHLDLGEALR 831
QY 157 SVVYDCFEAALGSEVRINWYTPRYEKELNHTQOLPALPMLNVSADGNVHLVNLVSEBQ 216
Db 832 ALVHR-RDTPLGT-----HTRTLP-VPHHRI---GSVVEALLLIA-RQ 868
QY 217 HFGLSLYNQVQGPCKPRWHKNLTGQIITLNHTDVLPCICIQVWPLEPDSVRTNICPR 276
Db 869 H-----GVPO-----LALTLEELVHAL-----D 886
QY 277 EDPAHQNLWQAARLRLTLQSWLLDAPCSLPAAEAL-----CWAPGGPCOPPLVPP 329
Db 887 RDPHS---WAAARL---LABALTRVPDATPTDVLRLADGIAERACDQGTQVQFQ 938
QY 330 LSWENVTVDKVLEFPPLKGNPCVQVNSSEKLOLQCLWADSLGPKDDVLLLETRGPQ 389
Db 939 AFWTAPRVPAATRLDLR-----RLVLAD-----GPP 965
QY 390 DNRSICALPESGCTSLPSKASTRAARLGEYLLQDLQSGOCLQ-LNDDDLGALMACPMKY 448
Db 966 H-----EPG-----PRHLDTAAG-----LIVADPTVQPLLVRFWDDERPLPATP 1005
QY 449 IHKRWALVWLACLLFAAALSILILLKKDHAKAAARAAAL-----LLYSADD 495
Db 1006 -----HATVATAQAALLHTRHRLDGLTEVLVDSTH 1037
QY 496 SGFERLVGALA-----SALCQLPLRVAVDLWSRRELSAOGPVAMFHAQR-----QT 542
Db 1038 RRADELLAVLAEEEPSALCR-----AVERWARDERPARHRAAVTHGLRTAPHARGADRT 1092
QY 543 LQEGGVVLLFSP-----GAVALCSEWLDQGVSGPCAHGPH-----DAFRA----- 583
Db 1093 LLRHAALVLLAGPSDPLRGGALALL-----VQDPCRDRHLPAALDLFAACDPYLLP 1145
QY 584 -SLSCVLPDLQGRAPGSYVGCACFDRLHHPDAVPALFR-----TVPVFTLPSQLPDLGAL 638
Db 1146 SAVAAALPHT-----PEPVLEAFRALLGCPDAGEALRLADATTPALT--HRVAALVGR 1198
QY 639 QQPRAPRSGRLOERAEQVSRL--OPALDYSFHPPT-----PAPGRGVGPGAGPGAGD 690
Db 1199 VTERPETAGHL---AAVYDRLDRDPAPRAVLLPLVTRLLDDGPEPARAA--LAGVLAAD 1253
QY 691 G 691
Db 1254 G 1254
```

## RESULT 3

H85840

hypothetical protein yehQ [imported] - Escherichia coli (strain O157:H7, substrain ED)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: H85840

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H85840  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-745 <STO>  
A:Cross-references: GB:AB005174; NID:gl2516346; PIDN:AA057188.1; GSPDB:GN00145; UMGF:232  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yehQ

Query Match 3.1%; Score 115.5; DB 2; Length 745;  
Best Local Similarity 21.4%; Pred. No. 0.83;  
Matches 162; Conservative 70; Mismatches 234; Indels 291; Gaps 43;

QY 21 LERLVGPQDATH-----CSPGLSCLWSDILCLPGDIVPAPGPVLAPTHLQTELVLRC 74  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 110 LEELATLPDATRKRAQVLVAKGITIELF-----CAPGEIPSAIRLP-MSDVRFYSSIRF 163  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 75 OKETDC---DCLRVAVHLAVGHWEPEDEKEF-----GGAADSGVEEPRNA 119  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 164 AR-CDCTEGTICEHV--LAVQAFVEAKTQQAETHLIWQMRSEHTVSSDDPFASEGNA 220  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 120 SLQ-----AQVLSFQA-----YPTARCVLLEVPAPALVQFGQSGSVV 159  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 221 CROYVQQLSQALWLGISQPLIHYEAFSAQQAERCNR--WVSESLRQLRASV--- 274  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 160 YDCFEAA-----LGSEVRIWSTOPTRYEKELNHTQOL-----PALPWLNVASDG-- 203  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 275 -DAFHARASHYHAGECLRLAALNSR-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 329  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 204 ----DNVHLV-LNV--SEQHFGLSLYNQVQPPKPRWKNLGPQIITLNNHTDLPV 254  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 330 GEAKLDHLRLVSLGMRWQDIEHYGLRIWF-----TDPDGTSLHLIS--- 371  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 255 CLCIQVWPLEPDSVRTNICPFREDPRAHONLWQAARLRLTLQSWLL-----DAPCS 306  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 372 ----RSPRSEQENSPAATRLFSFQAGALAGQIVSOAKRS 410  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 307 LPAAALCWRAAGDPQCPPLV--PLSWENVTVDKVLEFPL-----LKGHPNLC 353  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 411 ADGELLATR-----NRLSSVVPPLSPDAW-----RMSAPLRQPGIVALREYLRQPPSC 460  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 354 V-----QVNSSEKLOQECL--WADSLGPLKDDVLLLETRGPQDNRSCLALEFSGCTSLP 406  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 461 IRPLNQVDNLFILPVAECISLQWSSRQTL--DAQVISGEG-EDN--LTLT-----SLP 509  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 407 SKASTRAA--RLGEYLLDQSGOCL-----OLWDDDD----- 436  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 510 ASASAPYAVERMAA-LIQOTDDPVLVSGFVSFVQGLTLEPQVMTKTRAWALDAETAP 568  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 437 ----LGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKKD--HAKAAARGAALLL 490  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 569 VVASLPSASVLPVPSTAHO-----LLMRCQALLIQLLHNGWRYQEOQAIGOAELLA 619  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSSAOGPVANFHAQRQTLOEGGVV 550  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 620 NDLTAVGFYRLAHVL-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 648  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 551 LLFSPGAVALCSEWL-----ODGVSGPGAHG-----PHDAFR-----ASLSCVLPDLQ 594  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 649 M-----NNGVLLCEBOLPPLQOQGLNRRGPFGECEFCIELRDPDHRFRWKNKLEFLLP 702  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 595 GRAPGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQL 631  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 703 -----EHGPAFPAI-----VDCYTSPPTL 721  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 597 APGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQL 631  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
RESULT 4

C99995  
hypothetical protein Ecs2931 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: C99995  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: C99995  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-745 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA036354.1; PID:gl3362400; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: Ecs2931

Query Match 3.1%; Score 114.5; DB 2; Length 745;  
Best Local Similarity 21.6%; Pred. No. 0.99;  
Matches 163; Conservative 69; Mismatches 236; Indels 287; Gaps 43;

QY 21 LERLVGPQDATH-----CSPGLSCLWSDILCLPGDIVPAPGPVLAPTHLQTELVLRC 74  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 110 LEELATLPDATRKRAQVLVAKGITIELF-----CAPGEIPSAIRLP-MSDVRFYSSIRF 163  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 75 OKETDC---DCLRVAVHLAVGHWEPEDEKEF-----GGAADSGVEEPRNA 119  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 164 AR-CDCTEGTICEHV--LAVQAFVEAKTQQAETHLIWQMRSEHTVSSDDPFASEGNA 220  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 120 SLQ-----AQVLSFQA-----YPTARCVLLEVPAPALVQFGQSGSVV 159  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 221 CROYVQQLSQALWLGISQPLIHYEAFSAQQAERCNR--WVSESLRQLRASV--- 274  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 160 YDCFEAA-----LGSEVRIWSTOPTRYEKELNHTQOL-----PALPWLNVASDG-- 203  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 275 -DAFHARASHYHAGECLRLAALNSR-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 329  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 204 ----DNVHLV-LNV--SEQHFGLSLYNQVQPPKPRWKNLGPQIITLNNHTDLPV 254  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 330 GEAKLDHLRLVSLGMRWQDIEHYGLRIWF-----TDPDGTSLHLIS--- 371  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 255 CLCIQVWPLEPDSVRTNICPFREDPRAHONLWQAARLRLTLQSWLL-----DAPCS 306  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 372 ----RSPRSEQENSPAATRLFSFQAGALAGQIVSOAKRS 410  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 307 LPAAALCWRAAGDPQCPPLV--PLSWENVTVDKVLEFPL-----LKGHPNLC 353  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 411 ADGELLATR-----NRLSSVVPPLSPDAW-----RMSAPLRQPGIVALREYLRQPPSC 460  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 354 V-----QVNSSEKLOQECL--WADSLGPLKDDVLLLETRGPQDNRSCLALEFSGCTSLP 406  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 461 IRPLNQVDNLFILPVAECISLQWSSRQTL--DAQVISGEG-EDN--LTLT-----SLP 509  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 407 SKASTRAA--RLGEYLLDQSGOCL-----OLWDDDD----- 436  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 510 ASASAPYAVERMAA-LIQOTDDPVLVSGFVSFVQGLTLEPQVMTKTRAWALDAETAP 568  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 437 ----LGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKKD--HAKAAARGAALLL 490  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 569 VVASLPSASVLPVPSTAHO-----LLMRCQALLIQLLHNGWRYQEOQAIGOAELLA 619  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSSAOGPVANFHAQRQTLOEGGVV 550  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 620 NDLTAVGFYRLAHVL-----LNCQAEMARRDSVGEVPPVWRTVVGSGIA 648  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 551 LLFSPGAVALCSEWL-----ODGVSGPGAHGPHDAFRASLSCVLPD-----FLQGR 596  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 649 M-----NNGVLLCEBOLPPLQOQGLNRRGPF--PGECEFCIELR--LPDHRFRWKNKLEF--- 698  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
QY 597 APGSYVGACFDRLHLPDAVPALFRTVPVFTLPSQL 631  
||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
RESULT 4

R; White, O.; Eisen, J. A.; Heidelberg, J. F.; Hickey, E. K.; Peterson, J. D.; Dodson, R.

```

RESULT      8
T30987
telomerase-associated protein 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30987
C:Harrington, L.; McPhail, T.; Mar, V.; Zhou, W.; Oulton, R.; Bass, M.B.; A:
Science 275, 973-977, 1997
A:Title: A mammalian telomerase-associated protein.
A:Reference number: Z20952; MUID: 97172559
A:Accession: T30987
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2629 <HAR>
A:Cross-references: EMBL:U86137; NID:g1835783; PID:g1835784; PIDN:AAC53043.1
Query Match      2.9%; Score 108; DB 2; Length 2629;
Best Local Similarity 20.8%; Pred No. 16;
Matches 95; Conservative 52; Mismatches 160; Indels 150; Gaps

Oy 263 LEPSVTRNICPFREDPRAHQL-----WQARLRLLTLQSWLLDAPCSLPA 309
||| : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1585 LVPDLLGRLVLYASSKPKPANKLPAAADVAVFTFLRQQAASLLTQYPL--LLLOQAASOPE 1542
||| : : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Query Match	2.9%;	Score 107;	DB 2;	Length 719;	
Best Local Similarity	21.1%;	Pred. No. 3.6;			
Matches	142;	Conservative 75;	Mismatches 221;	Indels 234;	Gaps
Qy	99	PEDEK7CGGAADSVGEPRNASLQAQVVL5FSQAYPTARCVLLEVPQPAALVQFGSQSGVS	158		
Db	106	PEDWEIYGEVERGETDVLVSPERLNSVDFR-----DOVLPRLAATGGL	151		
Qy	159	VYDCEAALGSEVRIMSYT-OPRYEKELNHTQOLPA-LPWLNVASDGNVHLVLNVSEEQ	216		
Db	152	VVD--EAHCISD---WGDFRPDIYRRLTMAELPAGVPVLATTATA-NARVTDAVEQL	205		
Qy	217	HFGLSLYMNQVGGPPKPRWHKNTGPQIITLNHTDLVPCLCIQWPLEDSVSRFTNICPFR	276		
Db	206	GTG-----AGDALVLRG-----PLDRESLRGLGVLP	232		
Qy	277	EDPRAHQNLMOAERLRLTLQSNLLDAPCSLPAEALCWRAPGDDPCQPLVPPLSWENWT	336		
Db	233	D--AAHRLAWLGERLGE-----PCSGIITLTV-----AA	261		
Qy	337	VDKVLPEPLLKGHP--NLCVQVNSSEKLIQECLEWADSLGLP-----	376		
Db	262	AEEATAAFRLQRQYVASYTGKTENARLOAEEDLLANRVRKALVATSAHGFGDKPDLGV	321		

Db	103	LWAGQQLLWKGSKFSEHKVPKPRACNLNTVHTVNTSDTLLLTWSNPYPDPNLYLNHLTVAVN	162
Qy	94	GHWEPEDEEKFGGAADSGVEEPRNASLQAOVFLSFOAYPTARCVLLE---VQVPAALVQF	151
Db	163	-IWSENDPAD-----FLIY---NVTYLEPSLRATAASTLKS	193
Qy	152	QSGVSGVYVDCFEALGSEVRIRWYTOPRYEKELNHTQQLPALPWLVSADGDNVHLVLN	211
Db	194	GIS-----YRARVRAWAQCY-----NTTWSSEWSPSTKWHNSYREPEFEQHLLLG	236
Qy	212	VSEHQHFGLSL-----YNNQVGGPKPR-----WHKNLTGPOI	244
Db	237	VSVCIVITLAVCLLCYISITIKKEWMDQIPNPARSLVAIIHQDAQSGQWEKRSRGQEP	296
Qy	245	ITLNH-----TDLVPCLCIQWLPEDPSVRTNICPFEPDRAHQNLWQAARLLTLTLQSW	299
Db	297	AKCPHWKNCITKLPCF-----LEHNKRD-----EDPH-----KAA---KEMPFQGS	336
Qy	300	LLDAPCSLPAAALCWRAPEGDPQCPVPLPSLWENVTVDKVLEFPL-----LKGH	349
Db	337	GKSAMCPVEISKTVLW-----PESISVWRVCVELFAPVCEEEVEEEKG-382	
Qy	350	PNLCVQVNSSEK-----LQIQECLWADSLGPKLDDVLLLETRG---PODNRSICA	396
Db	383	-SFCASPESSRRDDFQEGREGIVARLTESLFDLLG-----EENGSGFCQDDMGESCL	432
Qy	397	LEPSGCTS-----LPSKASTRAARLGEVLLQDLQSGQCLQWLWDDDLGALWACPMDKYI	449
Db	433	LPPSGSTSAAHPWDEFFSAGPKAAPWGK-----EQPLHLEPSPASPSTQSPDN---481	
Qy	450	HKRWALVWLACLLFAAALSLLLLKDKHAKAAARGRAALLLYSADDSGFERLVGLASAL	509
Db	482	-----LTC-----TETPLVIAGNPAYRSFSNSLSQSP	508
Qy	510	COLPLRVAVDWLSRRELS-----AQGPVAFHQAQRQTLOEGGVV	550
Db	509	C--PRELGPDPLLARHLIEEVEPEMPCVPQLSEPTTVPQPEPETWEQTLRRNLVQHAAAA	566
Qy	551	LLFSP--GAVALCSEWLQDGVN-----GPGAHGPHDAFRASL--SCVLPDLELG-GRAP	598
Db	567	PVSAPTSGYQEFVHAVEQGGTQASAVVGLGPPGEAGYKAFSSLLASSAVSPEKCGFGASS	626
Qy	599	GSYVGACFDRLI-----HPDAVPALFRTVPVFT-----LPSQLPDFLGA-637	
Db	627	GEEGYKFPQDLIPCCPGDPAPVP-----VPLFTGLDPRRPSQSSHLPSSSPEHLGLE	681
Qy	638	-----LQOPRAPRS-----GRIQERAEQVSRALQAPALDS	666
Db	682	PGEKVEDMPKPLPQEQATDPLVDSLGSIGIVYSALTCHLCCHLKQCHQEDGGGTQPMVAS	741
Qy	667	-----YFHPDCTAPGRGVCPPGAGP	686
Db	742	PCCGCCGCCDRSSPTTTLRAPDPSPGGVP	770
RESULT 11			
A49724			
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type H precursor - human			
N:Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated			
C:Species: Homo sapiens			
C:Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 16-Jun-2000			
R:Matuzaki, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, J. Biol. Chem. 269, 2075-2081, 1994			
A:Title: Molecular cloning of a human transmembrane-type protein tyrosine			
A:Reference number: A49724; MUID:94124561			
A:Accession: A49724			
A:Molecule type: mRNA			
A:Residues: 1-1118 <MATO>			
A:Cross-references: GB:D15049; NID:g475003; PIDN:BAA03645.1; PID:g475004			
C:Genetics:			
A:Gene: GDB:PTPRH; SAP-1			
A:Cross-references: GDB:305504			







A;Molecule type: DNA  
A;Residues: 1-1487 <P>  
A;Cross-references: GB:AE001274: NID:g3264850: PIDN:AAC24673.1; PID:g2266918; GSPDB:GN00  
A;Experimental source: strain MHOM/IL/81/Friedlin  
C;Genetics:  
A;Gene: L1439.2  
A;Map position: 1

Query Match 2.88; Score 104; DB 2; Length 1487;  
Best Local Similarity 23.7%; Pred. No. 15;  
Matches 63; Conservative 19; Mismatches 112; Indels 72; Gaps 9;

QY 444 PMDKYIHKRWALVWLACILFAALSLIILLKKDHAKAAARGRAALLLYSADDSGFERLVG 503  
Db 19 PLDAQGH-----VCHARRAATAPATVVSNAEEVTEARGSCAPLHASREDT--RDRDS 68

QY 504 ALASALCQLPLRVAVDLMSRR-----ELSAQGFVAMFHAQRROT-----OEGGVVILL 552  
Db 69 PVAAPQRQOPALNAEDEWRWRASRPASSSSSSPSAODEAHSDRVVYEDGODGG----- 123

QY 553 FSPGVALCSEWLQDGVSGPGAGHPDAFRASLSCVLPDFLOGRAPGSYVGACFDRLLHP 612  
Db 124 --EEADRSEWRHS-----GADGASDDAEASSSTA-----LHP 155

QY 613 DAVPALFRTVPVFTLPSQLPDFLGALQOPRAPRGRGRLOERAQVSRALQPALDSYFHPPG 672  
Db 156 PSPQLHQOEASPLPLSLTPVATCTQAKPP-----STTAPPPGLESDTTPPPQ 203

QY 673 TPAPGRGVGP-----GAGPGAGDGT 692  
Db 204 VVSOGKGTDPKRLCTAVAPASGRAT 229

Search completed: May 19, 2002, 14:11:33  
Job time: 6401 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 12:24:03 ; Search time 43.31 Seconds  
(without alignments)  
390.268 Million cell updates/sec

Title: US-09-608-918-2  
Perfect score: 3716  
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGCGDGT 692

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326.5	8.8	617	4	US-09-188-930-303
2	174	4.7	866	2	US-08-620-694A-10
3	174	4.7	866	3	US-09-022-255-10
4	174	4.7	866	3	US-09-022-696-10
5	174	4.7	866	3	US-08-978-773-4
6	174	4.7	866	3	US-09-022-253-10
7	174	4.7	866	3	US-09-022-260-10
8	174	4.7	866	4	US-09-022-259-10
9	174	4.7	866	4	US-09-022-257-10
10	165	4.4	864	2	US-08-620-694A-2
11	165	4.4	864	3	US-09-022-255-2
12	165	4.4	864	3	US-09-022-696-2
13	165	4.4	864	3	US-08-978-773-2
14	165	4.4	864	3	US-09-022-253-2
15	165	4.4	864	3	US-09-022-260-2
16	165	4.4	864	4	US-09-022-259-2
17	165	4.4	864	4	US-09-022-257-2
18	162	4.4	330	4	US-09-188-930-125
19	116.5	3.1	1711	2	US-08-342-930-2
20	112.5	3.0	1208	4	US-09-463-702A-2
21	108	2.9	2629	2	US-08-751-189-4
22	108	2.9	2629	2	US-09-060-836-4
23	108	2.9	2629	4	US-09-184-445-4
24	107	2.9	4928	4	US-09-036-987A-5
25	107	2.9	4928	4	US-09-370-700-5
26	106	2.9	4439	3	US-08-897-236-23
27	103.5	2.8	3739	3	US-09-320-878-2

28	101.5	2.7	469	3	US-08-753-007A-8	Sequence 8, Appli
29	101.5	2.7	469	4	US-09-398-496-8	Sequence 8, Appli
30	100	2.7	954	2	US-08-749-169A-3	Sequence 3, Appli
31	100	2.7	954	2	US-09-130-032A-3	Sequence 3, Appli
32	99	2.7	3724	2	US-08-804-227C-10	Sequence 10, Appli
33	99	2.7	3724	2	US-08-804-198-4	Sequence 4, Appli
34	98	2.6	623	1	US-08-653-740-7	Sequence 7, Appli
35	98	2.6	623	2	US-09-073-594-7	Sequence 7, Appli
36	98	2.6	623	3	US-09-273-925-7	Sequence 7, Appli
37	98	2.6	2205	1	US-08-093-453B-2	Sequence 32, Appli
38	97.5	2.6	647	3	US-08-753-007A-32	Sequence 32, Appli
39	97.5	2.6	647	4	US-09-398-496-32	Sequence 33, Appli
40	97.5	2.6	3739	4	US-09-105-537-33	Sequence 6, Appli
41	97.5	2.6	11877	4	US-09-105-537-6	Sequence 5, Appli
42	97	2.6	1326	4	US-09-147-236-5	Sequence 4, Appli
43	96	2.6	405	2	US-08-483-151-4	Sequence 50, Appli
44	96	2.6	1843	4	US-09-413-814-50	Sequence 4, Appli
45	96	2.6	3729	2	US-08-804-227C-4	

## ALIGNMENTS

### RESULT 1

US-09-188-930-303  
; Sequence 303, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 303  
; LENGTH: 617  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-188-930-303

Query Match	8.8%	Score 326.5;	DB 4;	Length 617;
Best Local Similarity	26.7%	Pred. No. 2e+22;		
Matches 132;	Conservative 56;	Mismatches 184;	Indels 123;	Gaps 19;
QY	237	KNLTGQIITLHNHTDLVPCLCIQVMPLEPDSVRTNICPFREDPRAH-ONLWQAAR-----	290	
DB	183	KIVSGHGVDPDYELLPCMCIEASYLQEDTVRRKKCFQSWPEAYGDFWISIRFTDYS	242	
QY	291	-----LRLLTQSLWLLDAPCSLPAEALCWRAPGGDCQCPPLVPPLSWENVTYDKVLEPL	345	
DB	243	QHQNVMALTLR-----CPLKLEASLCWRQDPLTPCETL-----PNATAQSEGWYI	289	
QY	346	LKG---HNNLCVQVNSSEKQLQECLEWADSLGLKDDVLLLETRGPQDNRSLCALPESGC	402	
DB	290	LENVDLHPOLCFKF-SFENSSHVEC--PHQSGSLPSTVSMDTAQO-----LTLHFSSR	341	
QY	403	TSLPSKASTRAARLG-----EYLLQDQSGCQLQWDDDLGAL	440	
DB	342	TYATFSAWSDPGLGPDTPMPVYISQFGSVVPVTLIIIPFLROENCILVWRSVDHPA	401	
QY	441	W---ACPMCKYIHKRWALVWLACLLFAAALSILILIKDHAKAAARAAALLYSADDSG	497	
DB	402	WKHVLCPDD-----APYPTQLLR---SLGSGTRPVLLHHAADSEA	440	
QY	498	FERVLGALASALCQLPLRVA-----VDLWSRRELSAQGPVAMFHAQRQTLLQEGGVV	549	

Db 441 ORRLVGAEL-----LFTALGGGRDVIWDLWEGTHVARIGPLPWLWAARVAREQGTIV 495

QY 550 VLFSPGAVLCSWLDQVSGPAGHGDFAFRSLSCVLPDFLQGRAPSGVVGACFDRL 609

Db 496 LLLWN-----CAGPSTACSGDPQASLRTLL-----CAAPRPLLLAYFSRL 536

QY 610 LHPDAVPALFRTVPVFLPSQLPDLGALQ-OPRAPRSG-----RLQERAEQVS-RA 659

Db 537 CAGKIDIPRLPRALPRYLLRDLRDLRALDAQPATLASSWSHLGAKRCLKNRLEQCHLLE 596

QY 660 LOPALDSYFHPGTP 674

Db 597 LERAKDDYQGSTNSP 611

RESULT 2

US-08-620-694A-10

; Sequence 10, Application US/08620694A

; Patent No. 5869286

; GENERAL INFORMATION:

; APPLICANT: Yao, Zhengbin

; APPLICANT: Spriggs, Melanie

; APPLICANT: Fanslow, William

; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: Apple Operating System 7.5.5

; SOFTWARE: Microsoft Word for Apple, Version 6.0.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/620,694A

; FILING DATE: 21 MARCH 1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/620,694A

; FILING DATE: 23 MARCH 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Perkins, Patricia Anne

; REGISTRATION NUMBER: 34,695

; REFERENCE/DOCKET NUMBER: 2617-B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206)587-0430

; TELEFAX: (206)

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 866 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-620-694A-10

Query Match 4.7%; Score 174; DB 2; Length 866;

Best Local Similarity 21.6%; Pred. No. 1.1e-07;

Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLIT--GPQIITLN-----HT---DLVPCICIQWPLEPDS-----VRTN-- 271

Db 66 RNLTPSPKDLQQLHFAHTQGDLPFAHIE-WTLQTDASILYEGAELSVLQLNTNER 124

QY 272 IC----PFREDPRAHQLMWAARLRLITLQSWLLDAPCS-----LPAAALCWRAPEGD 331

Db 125 LCVRFEEFLSKLRHHRRWR-----FTFSHFVVDPDQYEVTVHHLPKP-----IPDGD 172

QY 322 PCOP-----LVPPLSWENVTVDKVLEPPLKG-----HPNLCOVQNSSEKLOLOECLWADSL 373

Db 173 PNHQSKNFLVPDCEHARMKVT-----PCMSGSLMDPNITVETLEAHQLRVSFTLWNES- 227

QY 374 GPLKDDVLLERWGPQDNRSLSALEPSGTS-----LPSKASTRAARLGEY-----LL 421

Db 228 ----THYQILLTSFPH-----MENHSCFEHMHHP-----APRPEEFHQRSNVTLTL 270

QY 422 QDLQS-----GQCQLQWDDDL--GALWACP-----MDKYIHKRWAL 455

Db 271 RNLKGCRRHQVQIQPFSSCL-----NDCLRHSATVSCPEMPDTPETIPDYMWLVY--WFI 325

QY 456 VWLACLLFAAALSLLILLK-----KDHAKAA-----ARGRAALLLYS 492

Db 326 TGISILLGVSVILLIYVCMTRWLAGPGESEKYSDDTKYTDGLPAADLIPPLPKPRKVVIIYS 385

QY 493 ADDSGFERLVGALAS--ALCOLPLRVAVDLWSRRELSAOGPVAFWFAHQRROTLOEGGVVYL 551

Db 386 ADHPLYVDVVLKFAPELLTACGTEVALDLEEQAISEAGVMTWVGKQKQEMVESNKIIV 445

QY 552 LFSGAVLCSWLDQVSGPGA-----HGP--HDAFRASLSVLPDFLQGRAPGSYV 602

Db 446 LCSRGTRA---KW--QALLRGAPVRLRCDHGKPVGDLFTAAMNMLPDFKRPACFGTV 500

QY 603 GACFDBLLHPDAVPALFRTVPVFTLPSQLPDLGALQOPRAPRSGRLQERAEQVSRALOP 662

Db 501 VCYFSEVSCDGDVDFLGAAPRYPLMDREFEYVFRIDQLEMFQPGRM-HRVGELSG---- 555

QY 663 ALDSYFHPGP 672

Db 556 ---DNVLRSPG 563

RESULT 3

US-09-022-255-10

; Sequence 10, Application US/09022255

; Patent No. 6072033

; GENERAL INFORMATION:

; APPLICANT: Yao, Zhengbin

; APPLICANT: Spriggs, Melanie

; APPLICANT: Fanslow, William

; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: Apple Operating System 7.5.5

; SOFTWARE: Microsoft Word for Apple, Version 6.0.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/022,255

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/620,694

; FILING DATE: 21 MARCH 1996

; APPLICATION NUMBER: USSN 08/538,765

; FILING DATE: 7 AUGUST 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/410,535

; FILING DATE: 23 MARCH 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Perkins, Patricia Anne



```
Db 501 VCYFSEVSCDGDVDFLFGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG----- 555
QY 663 ALDSYFHPFG 672
Db 556 --DNYLRSPG 563

RESULT 5
US-08-978-773-4
; Sequence 4, Application US/08978773
; Patent No. 6083906
; GENERAL INFORMATION:
; APPLICANT: Troutt, Anthony
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple PowerMacintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for PowerMacintosh, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,773
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/052,525
; FILING DATE: 27 NOVEMBER 1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2623-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-978-773-4

Query Match 4.7%; Score 174; DB 3; Length 866;
Best Local Similarity 21.0%; Pred. No. 1.le-07;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

QY 237 KNLITN---GQILITN---HT---DLVPCLCIQVWPLPDS-----VRTN-- 271
Db 66 NLTSPSPKDLQLQHLFAHTQQGDLFPVAHIE-WLTQDASILYLEGAELSVLQLTNTR 124
QY 272 IC---PFREDPRAHQLNQAARLRLTLQSWLLDAPCS-----LPAEAAALCWAPGGD 321
Db 125 LCVRFELSLRHHRRWR-----FTFSHFVVDPOEYEVTVVHLKP-----IPGD 172
QY 322 PCQP-----LVPPLSWENVTVDKVLEPPLKG-----HPNLCVQVNSSEKQLQECIWDLSL 373
Db 173 PNHQSKNFLVPDCEHARKVTT---PCMSGGSLWDNITVETLEAHQLRYSFTLWNES- 227
QY 374 GPLKDDVLLLETGRQPDNLSICALEPSCGNS-----LFSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTGFPH-----MENHSCFEHMHIP-----APRPEEFHQRSNVTILT 270
QY 422 QDLQS-----GQCILQLWDDDL--GALWACP-----MDKYIHKRWAL 455
```

```
Db 271 RNKGGCCRHQVQIQPFSSCL-----NDCLRHSAVSCPEMPDTPPEIPDPYPLWVY--WFI 325
QY 456 VWLACLLFAAALSILLK-----KDHAKAA-----ARGRAALLLYS 492
Db 326 TGISILLVGSVILLIVCMTWRLAGPGSEKYSDDTKYTDGLPAADLIPPLPKPKVWIIYS 385
QY 493 ADDSGFERLVGALAS--ALCOLPLRVAVDLWSRRELSAQGPVAVFHAQRROTLOEGGVVVL 551
Db 386 ADHPLYVDVVLKFAQFLLTACGTEVALDLLEQAISEAGVMTWVGKQKQEMVESNSKIIV 445
QY 552 LFSPPGAVALCSEWLQDGVSGPA-----HGP--HDAFRASLSCVLPDFLOQGRAPGSYV 602
Db 446 LCSRGTRA---KW--QALLGRGAPVRLRCDHGPVGDLEFTAANNMILPDFKRPACFTYV 500
QY 603 GACFDRLLHPDAVPALFRTVPVFTLPSOLPDFLGAQQPRAPSGRLQRAEVSRAIQP 662
Db 501 VCYFSEVSCDGDVDFLFGAAPRYPLMDRFEVYFRIQDLEMFQPGRM-HRVGELSG----- 555
QY 663 ALDSYFHPFG 672
Db 556 --DNYLRSPG 563

RESULT 6
US-09-022-253-10
; Sequence 10, Application US/09022253
; Patent No. 6096305
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694
; FILING DATE: 21-MARCH-1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 866 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```









APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/620,694  
FILING DATE: 21 MARCH 1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-255-2

Query Match 4.4%; Score 165; DB 3; Length 864;  
Best Local Similarity 19.8%; Pred. No. 8.2e-07;  
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;  
QY 192 PALPW-----LNVSDGDNVHLVNL-----VSEQHGLSLYWNQVQPPP--KPRW--HKNL 239  
Db 14 PALGWLILLNVLAPGRASPRLLDFPAPVCAOE--GLSC---RVKNSTCLDDSWIHPKNL 68  
QY 240 T--GPQITLNL-----HTDLVPCLCIQVWPLEPDS-----VRTN--IC- 273  
Db 69 TPSSPNYIINLSVSTQGEHLVPVLUHVE-WLQTDASILYLEGAELSVLQNLNTERLCV 127  
QY 274 --PFREDPRAHNLQAARLRLTLQSLWLDAPCSLPABE-----ALCWAPGDDP--- 322  
Db 128 KFOFLSMLQHHKRW-----FSESHFVVD-----PGQEYEVTVHHLPKPIPDGDNHK 176  
QY 323 -----CQPLVPLSWENVTVKLEPPLKLGHNLCVQVNSSEKQLQCECLWADSLG 374  
Db 177 SKIIFVPCDESKMKTTCSSVSSGLWD-----PNITVETLDTQHLRVDFTLWNEST- 228  
QY 375 PLKDDVLLLETRGPQDNRSICALPESGCTSL-----PSKASTRAARLGEYLLQDLSQG 428  
Db 229 PIQ-----VLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFH--- 274  
QY 429 CLQLWDDDLGALWAC-----PMDKYIHKR 452  
Db 275 -----WCCHHHVQVQPPFSSCLNCLRHAVTVPCPVISNTVTPKPVADYI-PL 321  
QY 453 WA---LVNWLACLLFAAALSLILLK-----RDHAKAAARGRA 486

Db 322 WYGLITLTAILLVGSVIVLIITCMTWRLSGADQEKHGDDSKINGILLPVADLTPTPPPLRPK 381  
QY 487 ALLLYSADDSGFERLVGALASAL-CQLPLRVAVDLMSRRELSAQGPVAFHQAQRROTLOE 545  
Db 382 WVIVISADHPLVVEVVLKFAQLITACGTEVALDILLLEEQVISEVGVMTWVSKQKQDMVES 441  
QY 546 GGVVYLLFSPG-----AVAL-CSEWLQDGVSGPCAHGPHDAFRASLSCVLP 590  
Db 442 NSKIIILCSRGTOAKWKAILGWAEPVQLRCDHWKPAQ-----DLFTAAMNMLP 491  
QY 591 DFLQGRAPGSYVACGDFRLLHDPALFRTVTVPTLPSQLPDLFGALQOQRPAPRGRLO 650  
Db 492 DFKRPACFGTYVVCYFSGICSERDVPDLFNITSRYPLMDRFEVYFRIQDLEMFEPGRMH 551  
QY 651 ERAEQVSRALQPALDSYFHP 671  
Db 552 HVRELTG-----DNYLOSP 565  
RESULT 12  
US-09-022-696-2  
Sequence 2, Application US/09022696  
Patent No. 6072037  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-696-2

Query Match 4.4%; Score 165; DB 3; Length 864;  
Best Local Similarity 19.8%; Pred. No. 8.2e-07;  
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;

QY 192 PALPW-----LNVSADGDNVHLVLN-----VSEHQHGLSLYNVOGPP--KPRW--HKNL 239  
Db 14 PALGWLILLNLVAPGRASPRLLDFPAPCAQE--GLSC---RVKNSTCLDSDSWHPKNL 68  
QY 240 T--GPOIITLN-----HTDLVPCICIQVWPLEPDS-----VRTN--IC- 273  
Db 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLQTDASILYLEGAELSVLQNTNERLCV 127  
QY 274 --PFREDPRAHQNWAARLLLTQSWLLDAPCSLPAEA-----ALCWAPGGDP--- 322  
Db 128 KFQFLSMLOHHRWR-----FSFSHFVVD-----PGQYEYVTVHHLKPIPDGDPNHK 176  
QY 323 -----COPLVPPLSWENVTVKVLFFLLKGNHPLCVQVNSSEKLOLQOECLWADSLG 374  
Db 177 SKIIFVPCEDSKMKMTTSCVSSGLWD-----PNTVETLDTQHLRVDFTLWNEST- 228  
QY 375 PLKDDVLLLTETRGPDNRSLCALEPSGCTSL-----PSKASTRAARLGEYLLODLOSQ 428  
Db 229 PYQ---VLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFN--- 274  
QY 429 CLQLWDDDLGALWAC-----WCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PL 321  
Db 275 -----WCCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PL 321  
QY 453 WA---LVWLACLLFAAALSILLK-----KDHAKAAARGRA 486  
Db 322 WYGLITLIALLVGSLVILICHTWRLSGADQEKHGDDSKINGILPVADLTTPPLRPKR 381  
QY 487 ALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAOGPVAFWHAQRRTLOE 545  
Db 382 WVIVSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTVWSRQKQEMVES 441  
QY 546 GGVVLLFSPG-----AVAL-CSEWLQDGVSGPGAHGPHDAFRASLSCLVLP 590  
Db 442 NSKIIILCSRGTOAKKAILGWAEPVQLRCDHWKPAG-----DLFTAAMNMILP 491  
QY 591 DFLQGRAPGSYVGACFDRLLHPDAVPALFRVTVFTLPSQLPDFLQALQOQPRAPSRGLQ 650  
Db 492 DFRPACFGTYVVCYFGSICSERVDPDLNITSRYPLMDRFEVYFRIQDLEMFEPGRMH 551  
QY 651 ERAEOVSRALQALDSYFHPP 671  
Db 552 HVRELTG-----DNVLOSP 565

## RESULT 13

US-08-978-773-2  
; Sequence 2, Application US/08978773  
; Patent No. 6083906  
; GENERAL INFORMATION:  
; APPLICANT: Troutt, Anthony  
; TITLE OF INVENTION: Method of Regulating Nitric Oxide Production  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple PowerMacintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft word for PowerMacintosh, Version. 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,773  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 60/052,525  
; FILING DATE: 27 NOVEMBER 1996

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 2623-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 864 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-978-773-2  
Query Match 4.4%; Score 165; DB 3; Length 864;  
Best Local Similarity 19.8%; Pred. No. 8.2e-07;  
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;  
QY 192 PALPW-----LNVSADGDNVHLVLN-----VSEHQHGLSLYNVOGPP--KPRW--HKNL 239  
Db 14 PALGWLILLNLVAPGRASPRLLDFPAPCAQE--GLSC---RVKNSTCLDSDSWHPKNL 68  
QY 240 T--GPOIITLN-----HTDLVPCICIQVWPLEPDS-----VRTN--IC- 273  
Db 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLQTDASILYLEGAELSVLQNTNERLCV 127  
QY 274 --PFREDPRAHQNWAARLLLTQSWLLDAPCSLPAEA-----ALCWAPGGDP--- 322  
Db 128 KFQFLSMLOHHRWR-----FSFSHFVVD-----PGQYEYVTVHHLKPIPDGDPNHK 176  
QY 323 -----COPLVPPLSWENVTVKVLFFLLKGNHPLCVQVNSSEKLOLQOECLWADSLG 374  
Db 177 SKIIFVPCEDSKMKMTTSCVSSGLWD-----PNTVETLDTQHLRVDFTLWNEST- 228  
QY 375 PLKDDVLLLTETRGPDNRSLCALEPSGCTSL-----PSKASTRAARLGEYLLODLOSQ 428  
Db 229 PYQ---VLESFSDSENHS-----CFDVVKQIFAPRQEEFHQRANVTFTLSKFN--- 274  
QY 429 CLQLWDDDLGALWAC-----WCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PL 321  
Db 275 -----WCCHHHVQVQPFSSCLNCLRHAVTVPCPVISNTTVPKPVADYI-PL 321  
QY 453 WA---LVWLACLLFAAALSILLK-----KDHAKAAARGRA 486  
Db 322 WYGLITLIALLVGSLVILICHTWRLSGADQEKHGDDSKINGILPVADLTTPPLRPKR 381  
QY 487 ALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAOGPVAFWHAQRRTLOE 545  
Db 382 WVIVSADHPLYVEVVLKFAQLITACGTEVALDLLEEQVISEVGVMTVWSRQKQEMVES 441  
QY 546 GGVVLLFSPG-----AVAL-CSEWLQDGVSGPGAHGPHDAFRASLSCLVLP 590  
Db 442 NSKIIILCSRGTOAKKAILGWAEPVQLRCDHWKPAG-----DLFTAAMNMILP 491  
QY 591 DFLQGRAPGSYVGACFDRLLHPDAVPALFRVTVFTLPSQLPDFLQALQOQPRAPSRGLQ 650  
Db 492 DFRPACFGTYVVCYFGSICSERVDPDLNITSRYPLMDRFEVYFRIQDLEMFEPGRMH 551  
QY 651 ERAEOVSRALQALDSYFHPP 671  
Db 552 HVRELTG-----DNVLOSP 565  
RESULT 14  
US-09-022-253-2  
; Sequence 2, Application US/0902253  
; Patent No. 6096305  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie

```

1  APPLICANT: Fanslow, William
2  TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17
3  NUMBER OF SEQUENCES: 10
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Immunex Corporation
6  STREET: 51 University Street
7  CITY: Seattle
8  STATE: WA
9  COUNTRY: USA
10 ZIP: 98101
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: Apple Power Macintosh
15 OPERATING SYSTEM: Apple Operating System 7.5.5
16 SOFTWARE: Microsoft Word for Apple, Version 6.0.1
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/09/022,253
19 FILING DATE:
20 CLASSIFICATION:
21
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US/08/620,694
24 FILING DATE: 21-MARCH-1996
25 APPLICATION NUMBER: USSN 08/538,765
26 FILING DATE: 7 AUGUST 1995
27 CLASSIFICATION:
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: USSN 08/410,535
31 FILING DATE: 23 MARCH 1995
32 CLASSIFICATION:
33
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Perkins, Patricia Anne
36 REGISTRATION NUMBER: 34,695
37 REFERENCE/DOCKET NUMBER: 2617-B
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: (206)587-0430
40 TELEFAX: (206)
41
42 INFORMATION FOR SEQ ID NO: 2:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 864 amino acids
45 TYPE: amino acid
46 TOPOLOGY: linear
47 MOLECULE TYPE: protein
48
49 US-09-022-253-2

```

Query Match	4.4%;	Score 165;	DB 3;	Length 864;
Best Local Similarity	19.8%;	Pred. No. 8.2e-07;		
Matches 123;	Conservative 79;	Mismatches 209;	Indels 210;	Gaps 31;
QY	192	PALPW-----UNVSADGDNVHLVLN-----VSEEQHFGLSYWNQVGGPP--KPRW--HKNL	239	
Db	14	PALGWLLLLLVNVLAPGRASPRLLDFPAPVCAQE--GLSC--RVKNSTCLDDSWTHPKNL	68	
QY	240	T--GPQIITLN-----HTDLVPCICIQWVPLEPDS-----VRTN--TC-	273	
Db	69	TPSSPKNIYINLSVSSQHGELVPVLHVE-WTLQTDASILYLEGAELSVLQLTNERLCV	127	
QY	274	--FPREDPRAQNWLQAARLRLTLQSWLLDAPCSLPAEA-----ALCWRAPOGDP---	322	
Db	128	KFOFLSMQLQHRKWR-----ESFSHFVYD-----PGQEYEVTVHHLPKPIPDGDPNHK	176	
QY	323	-----CQPLVPPLSWENVVDKVFLLKGHPNLQVQVNSSEKQLQECUWADSLG	374	
Db	177	SKIIFVPDCEBSRKMKTTSVSSGSLWD-----PNITVETLDTQHLRVDFTLWNEST-	228	
QY	375	PLKDDVLLTETRGPDNRSCLALEPSCSTSL-----PSKASTRAARLGEYLLQDLOGSQ	428	
Db	229	PYO-----VLLESFSDSENHS-----CFDVPVKQIFAPROEEFHQRANVTFTLSKEH---	274	
QY	429	CLQLWDDDLGALWAC-----PMDKYIHKR	452	
Db	275	-----MCCHHHVQVQPFSSCLNDCLRHAVTVPCPVISNTNTPVKPVADYI-PL	321	

```

QY 453 WA---LWLACLFAAALSLLILLK-----KDHAKAAARGRA 48
DB 322 WYGLITLAIILVGSVIVLITICMTWRLSGADQEKHGDSKINGILPVDLTPPLRPRK 381
QY 487 ALLIYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRRLSAGQPVAVFHAOROTLQOE 545
DB 382 VMIYVSDHPDIYEVVLKFAQFLITACGTEVALDLLEEQVISEVGVMTWVSROKQEMVES 441
QY 546 GGVVVILFSFG-----AVAL-CSEWLODGVSGPGAHPDHFDAFRASLSQVLP 590
DB 442 NSKIILCSRCTQAKKAILGWAEPVQLRCDHKPKAG-----DLFTAAMNWLIP 491
QY 591 DFLQGRAPGSYVGACFDBRLHPDPAVPALFRTVPVFTLPSQLPDLGALQOPRAPSRLQOE 650
DB 492 DFKPACPGFTYVVCYFSGICSERVDPLFNITSRYPLMDRFEVYFRIQDLEMFEFGRMH 551
QY 651 ERAEQVSRALQPALDSYFHP 671
DB 552 HVRELTG-----DNYLOSP 565

RESULT 15
US-09-022-260-2
; Sequence 2, Application US/09022260
; Patent No. 6100235
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,260
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/620,694
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 864 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-022-260-2

```

Query Match 4.4%; Score 165; DB 3; Length 864;  
Best Local Similarity 19.8%; Pred. No. 8.2e-07;

```
Matches 123; Conservative 79; Mismatches 209; Indels 210; Gaps 31;
QY 192 PALPW----LNVSDGDNVHLVN----VSEQHFGSLSYWNOVGPP--KPRW--HKNL 239
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 14 PALGWLILLNVLAPGRASPRLDPPAPCAQE--GLSC---RVKNSTCLDDSWIHPKNL 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 240 T--GQIITLN-----HTDLVPCICIQWPLEPDS-----VRTN--IC- 273
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 TPSSPKNIYINLSVSTQHGLVPLVHVE-WTLQTDASILYLEGAELSVLQNTNERLCV 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 274 --PFREDPRAHONLWQARLRLTLQSWLLDAPCSLPABE-----ALCWRAPEGDP--- 322
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 128 KFQFLSMLQHHKRW-----FSFSEFVVD-----PGQEYEVTVHHLKPIPDGDPNHK 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 323 -----CQPLVPPLSWENVTVDKVLFPILKGNLVCQVNSSEKLOLQECLEWADSLG 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 177 SKIIFVPCEDSKMKMTTSCVSSGSLMD-----PNITVETLDTQHLRVDFTLWNEST- 228
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 375 PLKDDVLLLETRGPDNRSICALPFGCTSL-----PSKASTRAARLGEYLLLODLQSGQ 428
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 229 PYQ-----VLEFSFSENHS-----CFDVVKQIFAPRQEEPHQRANVTFTLSKEH--- 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 429 CLQLWDDDLGALWAC-----WCCHHHVQVQPFSSCLNDCLRHAVTVPCPVISNTTVKPVADYI-PL 321
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 275 -----LWNLACLLFAAALSILILLK-----KDHAKEAAARGRA 486
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 453 WA---LWNLACLLFAAALSILILLK-----KDHAKEAAARGRA 486
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 322 WYGLITLITAILLVGSIIVIIICMTWRLSGADQKHGDDSKINGILFVADLTTPPLRPK 381
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 487 ALLYSADDSGFERLVGALASAL-CQLPLRVAVDLWSRRELSAQGPVAFHQRROTLOE 545
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 382 VMIVYSADHPLVEVVLKFAQFLITACGTEVALDLLEEQVISEVGVNTWWSRQKEWVES 441
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 546 GGVVVLLFSPG-----AVAL-CSEWLQDYGSGPGAHGPHDAFRASLSCLVP 590
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 442 NSKIIILCSRGTOAKWKAILGHAEPVQLRCDHWKPAG-----DLFTAMNWILP 491
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 591 DFLQGRAPGSGYVACFDRLHDPDPAVPALPRTVPVFTLPQLPDLGALQOQPRAPRSRLQ 650
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 492 DFKRPACFGTYVVCYFSGICSERVDPDLFNITSRYPLMDRFEVYFRIQDLENFEGRMH 551
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 651 ERAEQVSRALQPALDSYFHP 671
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 552 HVRELTG-----DNYLQSP 565
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 19, 2002, 14:10:14  
Job time: 6371 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 13:26:44 ; Search time 115.29 seconds  
(without alignments)  
1038.360 Million cell updates/sec

Title: US-09-608-918-2  
Perfect score: 3716  
Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPGRGVGPGAGGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertibrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2598	69.9	538	4 Q9BR97	Q9br97 homo sapien
2	1812.5	48.8	567	11 Q9J43	Q9j43 mus musculu
3	174	4.7	866	4 O43844	O43844 homo sapien
4	173	4.7	866	4 Q96F46	Q96f46 homo sapien
5	165	4.4	864	11 Q60943	Q60943 mus musculu
6	131.5	3.5	4340	2 Q30764	Q30764 streptomyc
7	130.5	3.5	1366	2 Q92529	Q92529 streptomyc
8	116	3.1	1021	4 Q96M86	Q96m86 homo sapien
9	114.5	3.1	745	2 Q9EYF6	Q9eyf6 escherichia
10	114.5	3.1	1193	16 Q9HZE5	Q9hze5 pseudomonas
11	112.5	3.0	657	4 Q96DW2	Q96dw2 homo sapien
12	112.5	3.0	744	4 Q96F55	Q96f55 homo sapien
13	111.5	3.0	478	16 Q9RY89	Q9ry89 deinococcus
14	111	3.0	1279	12 Q66031	Q66031 cercopithec
15	110	3.0	757	6 Q9N015	Q9n015 macaca fasc
16	109.5	2.9	620	4 Q9HCN6	Q9hcn6 homo sapien

17	109	2.9	427	2 Q9SIL6	Q9sll6 streptovirt
18	109	2.9	603	12 P89140	P89140 pseudorabie
19	109	2.9	604	12 Q9YVB7	Q9yvb7 pseudorabie
20	109	2.9	1343	12 Q06635	Q06635 bovine herp
21	108.5	2.9	1072	4 Q96EP0	Q96ep0 homo sapien
22	108	2.9	361	2 O86440	O86440 pseudomonas
23	108	2.9	2629	11 P97499	P97499 mus musculu
24	108	2.9	10917	2 Q93NW6	Q93nw6 streptomyc
25	107	2.9	719	2 O69992	O69992 streptomyc
26	107	2.9	825	4 Q96P01	Q96p01 homo sapien
27	107	2.9	1001	2 Q9KZM8	Q9kzm8 streptomyc
28	107	2.9	1118	4 Q15426	Q15426 homo sapien
29	107	2.9	2003	4 O00306	O00306 homo sapien
30	107	2.9	4928	2 Q9ALM3	Q9alm3 saccharopol
31	106.5	2.9	679	2 Q9F2P6	Q9f2p6 streptomyc
32	106	2.9	1211	11 Q9R0L3	Q9r0l3 rattus norv
33	106	2.9	1235	11 Q9R0L5	Q9r0l5 rattus norv
34	106	2.9	1273	11 Q9R0L4	Q9r0l4 rattus norv
35	106	2.9	1531	5 Q9W0P8	Q9w0p8 drosophila
36	106	2.9	9510	2 Q93NX9	Q93nx9 streptomyc
37	105.5	2.8	1354	11 Q9EPW8	Q9epw8 mus musculu
38	105.5	2.8	1542	4 O15035	O15035 homo sapien
39	105	2.8	1766	12 Q9J599	Q9j599 fowlpox vir
40	104.5	2.8	238	16 O53500	O53500 mycobacteri
41	104	2.8	1256	11 Q99W76	Q99w76 rattus norv
42	104	2.8	1487	5 O15843	O15843 leishmania
43	103.5	2.8	2055	4 O75055	O75055 homo sapien
44	103	2.8	414	4 Q9H612	Q9h612 homo sapien
45	103	2.8	643	12 Q85050	Q85050 pseudorabie

ALIGNMENTS

RESULT 1

Q9BR97	PRELIMINARY;	PRT;	538 AA.
ID	Q9BR97		
AC	Q9BR97;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	HYPOTHETICAL 59.1 KDA PROTEIN.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-ENDOMETRIAL ADENOCARCINOMA;		
RA	Strausberg R.;		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC006411; AAH06411.1; -		
KW	Hypothetical protein.		
SQ	SEQUENCE 538 AA; 59127 MW; 65E6344DA6A5AFD2 CRC64;		

Query Match 69.9%; Score 2598; DB 4; Length 538;  
Best Local Similarity 99.4%; Pred. No. 8e-208;  
Matches 480; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MPVPWFLLSLALGRSPVLSRLERLVGPDATHCSFGLSCRLWDSIDLCPGDIVPARGPV	60
Db	1	MPVPWFLLSLALGRSPVLSRLERLVGPDATHCSFGLSCRLWDSIDLCPGDIVPARGPV	60
Qy	61	LAPHLQTELVLCQKTCDDCLRLVAVHVLAVHGHWEPEDEEFKGAADSGVEPRNAS	120
Db	61	LAPHLQTELVLCQKTCDDCLRLVAVHVLAVHGHWEPEDEEFKGAADSGVEPRNAS	120
Qy	121	LQAGVLSFOAYPTARCVCVLEVVQVPAALVQFGSGSVVYDCFEAALGSEVRIMSYQPR	180
Db	121	LQAGVLSFOAYPTARCVCVLEVVQVPAALVQFGSGSVVYDCFEAALGSEVRIMSYQPR	180
Qy	181	YEKELNHTOOLPALPWLNVNSADGDNVHLVNLNVSEQHFGLSLYWNQVQGPCKPRHKNLT	240

```
Db 181 YEKLNHTQOLPALPNLWNSADGDNVHLVNLVSEQHFGLSLYWNQVGGPKPRWHKNT 240
Qy 241 GPOIITLNTDLVPCLCIOWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLTQSWL 300
Db 241 GPOIITLNTDLVPCLCIOWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLTQSWL 300
Qy 301 LDAPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEFPLKKGHPNLCVQVNSSE 360
Db 301 LDAPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEFPLKKGHPNLCVQVNSSE 360
Qy 361 KLQQLCEWLWADSLGPKLDDVLLLETRGPQDNRSICALPEPSGCTSLPSKASTRAARLGEYL 420
Db 361 KLQQLCEWLWADSLGPKLDDVLLLETRGPQDNRSICALPEPSGCTSLPSKASTRAARLGEYL 420
Qy 421 LDQLOSQCQLQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALLSLTLKKDHAKA 480
Db 421 LDQLOSQCQLQLWDDDLGALWACPMCKYTHKRWALVWLACLLFAAALLSLTLKKDHAKA 480
Qy 481 AAR 483
Db 481 WLR 483

RESULT 2
ID Q99J43 PRELIMINARY; PRT; 567 AA.
AC Q99J43;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE HYPOTHETICAL 62.8 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004759; AAH04759.1;
KW Hypothetical protein.
SQ SEQUENCE 567 AA; 62798 MW; C1AAAB79E2006B1D CRC64;
```

```
Query Match 48.8%; Score 1812.5; DB 11; Length 567;
Best Local Similarity 62.9%; Pred. No. 21e-142;
Matches 357; Conservative 50; Mismatches 130; Indels 31; Gaps 7;

Qy 1 MPVPFLLSLALGRSPVVLSERLVGPQDATHCSGLSCLRLWDSIDLCLPGDIVPAPGV 60
Db 1 MPVSNFLSLALGRNPVVVSLERLMEPQDARCSGLSCHLWDGDLCLFGLSLOSAPGV 60

Qy 61 LAPTHQLTELVRCKETDCDLCRLVAVHLVHGWEEPEDEKFGGAADSGVEEPRNAS 120
Db 61 LVPTRLQTELVRCKQKTDLCALCVRVVHLVHGWAEPEE-----AGKSDSELQESKNAS 116

Qy 121 LQAQVVLFSQAYPTARCVLLEVOVPAALVQFGSGSVVYDCFEAALGSEVRWTSYQPR 180
Db 117 LQAQVVLFSQAYPIARCALLEVOVPAALVQFGSGSVVYDCFEAALGSEVRWTSYQPR 176

Qy 181 YEKLNHTQOLP-----ALPVLNVSADGDNVHLVNLVSEQHFGLSLYWN 225
Db 177 YQKELNLTQOLPDCRGLEVRDYSIQSGWLPVLNVLSTGDGNVLTLDVSEQDFSLYLRL 236

Qy 226 VQGGPKPRWHKNTLGPQIITLNTDLVPCLCIOWPLEPDSVRTNICPFREDPRAHONL 285
Db 237 PVPDALKSLWYKNTLGPQIITLNTDLVPCLCIOWPLEPDSVRTNICPFREDPRAHONL 296

Qy 286 WQAARLLTLTQSWLDDPCSLPAEALCWAPGDCQPLVPLSWENVTVDKYLEFPL 345
Db 297 WHIARLRLSPGVWQLDAPCCLPKGVTLQWQAPDQSPCQPLVPPVPQKNATVNEPDQFL 356
```

```
Qy 346 LKGNHNLVQVNSSEKQLQCEWLWADSLGPKLDDVLLLETRGPQDNRSICALPEPSGCTSL 405
Db 357 VAGHPNLCVQVNSSEKVLQACWLWADSLGPKLDDVLLLETRGPQDNRSICALPEPSGCTPL 416
Qy 406 PSKASTRAARLGEYLQLOSQCQLW- DDDLGALWACPMCKYTHKRWALVWLACLLFA 464
Db 417 PSMASTRAARLGEELQDQFRSHOCQOLWDDNMGSLWACPMCKYTHRRWVWLWACLLFA 476
Qy 465 AALSILILLKKDHAKAAAGRAALLYS-----ADDSGFERLVGALASALCOLPLRVAV 518
Db 477 AALFFFLLLKKDRTWTPSPGSAACYPISCKAGRPAATSG-STSTGCGCTQTLC--PPRSAS 533
Qy 519 DLWSRRELQAQGPVAFWHAQRRTLQEG 546
Db 534 PRSSPCPRSCR--LSWMHCREAAPLPRG 559

RESULT 3
ID O43844 PRELIMINARY; PRT; 866 AA.
AC O43844;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE IL-17 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98035683; PubMed=9367539;
RA Yao Z., Spriggs M.K., Derry J.M.J., Strockbine L., Park L.S.,
RA VandenBos T., Zappone J., Painter S.L., Armitage R.J.;
RT "Molecular characterization of the human interleukin (IL)-17
RT receptor."
RL Cytokine 9:794-800(1997).
DR EMBL; U58917; AAB99730.1;
KW Receptor.
SQ SEQUENCE 866 AA; 96122 MW; 88AF626A83F3FF70 CRC64;
```

```
Query Match 4.7%; Score 174; DB 4; Length 866;
Best Local Similarity 21.6%; Pred. No. 8.6e-06;
Matches 119; Conservative 76; Mismatches 189; Indels 166; Gaps 32;

Qy 237 KNLIT--GPOIITLN-----HT---DLVPCLCIOWPLEPDS-----VRLT-- 271
Db 66 RNLTPSSPKDLQIQLHFAHTQOGLFPVAHIE-WTLQTDASILYLEGAELSVLQLTNTER 124

Qy 272 IC---PFREDPRAHONLWQAARLLTLTQSWLWDAPCS-----LPAAALCWAPGCD 321
Db 125 LCVRFELSLRHHRRWR-----FTFSHEVVDQDQYEVTVHHLPRK-----IPGD 172

Qy 322 PCQP---LVPLSWENVTVDKLEFFLLKG---HPNLCVQVNSSEKQLQCEWLWADSL 373
Db 173 PNHOSKNFLVPDCEHARKVTT---PCMSSGLWDPNITVETLEAHLQVSVFLWNES- 227

Qy 374 GPLKDDVLLLETRGPQDNRSICALPEPSGCTSL---LPSKASTRAARLGEY-----LL 421
Db 228 ---THYQILLTSFPH-----MENHSCFEHMHIP-----APRPEFHORSNVTLTL 270

Qy 422 QDLOS-----GQCQLWDDDL--GALWACP-----MDKYTHKRWAL 455
Db 271 RNLKGCGRHQVQIQPFSSCL---NDCLRHSATVSCPEMDPTPEPIQDMPPLWY--WFI 325

Qy 456 VWLACLLFAAALLSLTLK-----KDHAKEA-----ARGAALLYS 492
Db 326 TGISILVGVSVLLIVCTWRLAGPSEKYSDDTKYTDGLPAADLIPLPKPRKWIYS 385

Qy 493 ADDSGFERLVGALAS--ALCOLPLRVAVDLWSRRELQAQGPVAFWHAQRRTLQEGVVVL 551
Db 386 ADHPLYVDVVLKFAQLLTACGTGTEVALDLEEQAISEAGVMTWVGRQKQENVESNKIIV 445
```

OY	552	LFSGAAALCSEWJLDDGVSGGA-----HGP--HDAFRASISCVLPDLQGRAGSSV	602
Dd	446	LCSSKSTRA--KM--QAIIIGAGATRYLRCDHGKVGGLFTRAMMILPDKRPACTETV	500
OY	603	GACFDRLHPDVAVALFRTVEVFVLPSQLDPFLGALOOPRAPRSGRIQERAEQVSRALQP	662
Dd	501	VCFSESVSCSDGDVDPLDFGAPARYPLDMRFEEVYPRIQDLEMFQGRM-HRVGELSG----	555
OY	663	ALDSYFHHPG 672	
Dd	556	--DNITLNSPG 563	
 RESULT 4			
ID	O96FA6	PRELIMINARY;	PRT; 866 AA.
AD	O96FA6;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	SIMILAR TO INTERLEUKIN 17 RECEPTOR.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=UTERUS, AND LEIOMYOSARCOMA;		
RA	Strausberg R.;		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL, BC011624, AHH11624.1; --		
KW	Receptor.		
SO	Sequence	866 AA;	96131 MW; 28330BED2303BDC9 CRC64;

```

Db      501 VCYFEVSCDDGDDPLFAAPRYPFLMDRFFEEVYRIDLDFMGGRM-HRVGELSG----- 555
Oy      663 ALDSTFHPGP 672
          1:1
          11
Db      556 --DNYLRSPG 563

RESULT      5
Q60943      Q60943      PRELIMINARY;      PRT;      864 AA.
AC          Q60943:
DT          01-NOV-1996 (TREMBLrel. 01, Created)
DT          01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT          01-DEC-2001 (TREMBLrel. 15, Last annotation update)
DE          INTERLEUKIN I7 RECEPTOR.
GN          IL17R.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_Taxid=10090;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          TISSUE=THYMOMA EL4;
RX          MEDLINE=96111968; PubMed=8777726;
RA          Yao Z., Fanslow W.C., Seldin M.F., Rousseau A.M., Painter S.L.,
RA          Comeau M.R., Cohen J.I., Spriggs W.K.;
RT          "Hepesvirus Salmir1 encodes a new cytokine, IL-17, which binds to a
RT          novel cytokine receptor."
RL          Immunity 3:811-821(1995).
DR          EMBL: U31993; AAC2357.1; -.
DR          MGI:107399; I117r.
SQ      RECEPTOR.      864 AA;      97807 MW;      3435D51AA687DA31 CRC64;

```

```

Query Match          4.7%; Score 173; DB 4; Length 866;
Best Local Similarity 21.3%; Pred. No. 1e-05;
Matches 117; Conservative 76; Mismatches 191; Indels 166; Gaps 31

OY 237 KNLTL--GPOILTYLN---HT---DLVPCICIOVMPLRPS-----VETN-- 271
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 66 KMLTSSPKRDIOQLHFATITQOGDLFPVANHIE-WTLQTSALILYEGALSLVQLWTNR 124
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 272 IC---PREDPRAHONLWQAARLRLTLQSWLLDAPCS-----LPAEALCWRAPGD 321
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 125 LCVREFEFLSKLHHNHRMR-----FTSFHVDPDQEEYEVTHHLRP-----IPGD 172
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 322 PCQP---LVPRLSMENYTVDKYLEPRLKG---HPNLCQVNSSEKIQLOECMLADSL 373
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 173 PPHQSKNFIIVPCSEHARMKVTT---PCMSGSLMDPNTVTELEHNOIRVSEFTLWNE- 227
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 374 GFLKEDVLLLETRGHODNSLCALERPSCGTS---LPSKASTRAARLGEY-----TL 421
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 228 ---THYQILTLSPFH---MEHNSCFEIMNHLP-----APRPEEFHQRSNVTLTL 270
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 422 ODLOS-----GQCQLMDDDL--GALMACP-----MDKYHKKRMAL 455
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 271 RMLKGCGRHQVQIOPEFFSSCL---NDCLRHSAIVSCPEMRDPREPIIDYMLPWV--WPI 325
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 456 VWLACLIFPAALSLILLK-----KDAKKAARGRALLYS 492
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 326 TGISILSLVSVTLILYICMTMRLAGPGESEKYSDTKTYDGLPYADLIRPLPKRKWIIYS 385
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 493 ADDSGFERLYGALAS--ALQGLPLRYAVADMSRELISAOGPVAMFNOROTLOEGSVVVL 551
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 386 ADHPLYVDVNLKFAOFLPLRACGT EYALDLLEDAIS EAGVMTVWGROKOMEVNSKTIIV 445
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 552 LFPSPGAVALCSMLODGVSGPGA-----HGP--HDAERASTLSCVLPFLQGRAGSIV 602
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 446 LCSRGRRA---RW--QALLGRGAPVRLRCDHKGKRPVQDLFTTAAMNMILTRPFKRRACGTIV 500
      :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
OY 603 GACFDRLLHPDAVPALEFRTVPVFTLPSQLPDFLGALQDPRAPRSGRLQERAEVSAIAP 662

```

Query Match	4.4%	Score 165	DB 11	Length 864
Best Local Similarity	19.8%	Pred. No. 4.8e-05		
Matches 123	Conservative 79	Mismatches 209	Indels 210	Gaps 31
QY	192 PALPW-----LANSAGDNVHLVNI-----VSEOHFELSLYMNVOGPP--KPRW--HKNU	239		
Db	14 PALGMLLLLLNLNVLAFGRASPRLLDPAPRCAD--GLSC--RYKNSCTLDDSWIHPRKL	68		
QY	240 T--GPOITLN-----HTDLVPLCLIOVWPLEPDS-----VRTN--IC	273		
Db	69 TPSSPRNIYINLSVSTGELGVLPVLAHE-WTILQTDASILYLEGAEISVLQNTNERLCV	127		
QY	274 --PFREDPRAHONLMOANRLRLTLQSLMLDAPCSIPAA-----ALCMRAPGQDP--	322		
Db	128 KFOFLSMLOHKKRRMK-----FSESHPFVD-----PGGEYEVYVNIHKPRIPDGPNIHK	176		
QY	323 -----CQPLVPRPSMENVTVDKLEPRLLKGNHLGVQVNSSEKLOLOCLMAVDSIG	374		
Db	177 SKIIFVPRDCEDSKMKMTTSCVSSGSLMD-----PNTIVEPIIDQHLVADFTLIWNEST-	228		
QY	375 PLKDDVLLLETRGPODNRSLCALPESGCTSL-----PSKASTRAARLGEYLLDLOLQSQ	428		
Db	229 PLYQ-----VLLESFSDSENHS-----CDVYKQIAPRQOEHTQORANVTFLSKFH---	274		
QY	429 CLQLMDDLGALIMAC-----PMDKIYIKR	452		
Db	275 -----WCSHNHVQVQPFESSCLNDCLRNAVTVCPIYNTITVPRKVADYI--PL	321		
QY	453 MA---LWMLACLLPFAALSLILLLK-----KDHAKKAARGRA	486		
Db	322 WYVGLTLLAILLVGSVIVILITCMWRISGADQEKHGDSTKINGILPVADLTGPPILPRPK	381		
QY	487 ALLIYSADSGCFERLYGALASAL--COLPLRVAVVDLMSRRELSNOCGVAFHQAOROTLOE	545		
Db	382 WNIIVSADHPLYVEVYLKRAQFLITRACGTEVALLDLBEQVIVISGVCMIVVNSKQKQMVES	441		
QY	546 GGVVVLLFSPG-----AVAL--CSEWLIQDGVSGPAGHNPDAFRASLSCLVP	590		

Query Match 3.5%; Score 130.5; DB 2; Length 1366;  
Best Local Similarity 22.4%; Pred. No. 0.065;  
Matches 175: Conservative 57; Mismatches 238; Indels 311; Gaps 42;



QY 21 LERLVGPDATHCSPGLSCLRWSDILCLPGDIVPAPGPVLAPTHLOTELVLRCQKETDC 80  
 Db 675 LDGALADPDAGH---PLTIRLSEVRAALPGP--PAPVPVTRDAVFTAYLDL----- 721  
 QY 81 DLCLRVAVHLAVHGWEEPEDEEKGAA-----DSGVPEPRNASL-QAQQVLSFOA 131  
 Db 722 -MCLRVATRLA-----DENGIRGTAVRELAARKVSCQVHEAARRSLGPGCGGLDRES 771  
 QY 132 YPT-----ARCYLE-VQVPAAL-----VQFGQSVG 156  
 Db 772 FETLPGCPAPARLGGGWPAPVLAEGLFVPTGSGYRFAHEELADWIGTHLDLGEALR 831  
 QY 157 SVVYDCEAALGSEVRINSYTPRYEKELNHTQQLPALPWLNVSDGDNVHLVNSEQ 216  
 Db 832 ALVHR-RDTPGLT-----HTRTLP-VPHHRI---GSVVEALLLLA-RQ 868  
 QY 217 HFGLSLYWNQVQPKPRWHKNTLGPQITLNTDLPCLCIQVWPLEDPSVRTNICPFR 276  
 Db 869 H-----GVPO-----LALTLELVHAL-----D 886  
 QY 277 EDPRAHONLMOAARLRLTLQSWLDAPCSLPAEAL-----CWRAFGDPQCPPLVPP 329  
 Db 887 RDPHS-----WNAARL---LAEALTRVPDATPTDVLRLADGIAERAGDGQPTQVFGP 938  
 QY 330 LSWENVTVDKVLEPPLKGHENLQVQVNSSEKILQOECLWADSLGPKDDVLLLETRGQ 389  
 Db 939 AFWTAPRVPAATRLDRLR-----RLVLAD-----GPP 965  
 QY 390 DNRSLCALEPSGCTSLPSKASTRAARLGEYLLODQSGOCLOL-WDDDLGALWACPMKY 448  
 Db 966 H-----EPG-----PHLDTAG-----LIVADPTVQPLVRFWDDERPLPATP----- 1005  
 QY 449 IHKRWLVWLACLLFAAALSLLKDKHAKAAARGAAL-----LLYSADD 495  
 Db 1006 -----HATVATAQAALLHHRGLDGLTEVLVDSTH 1037  
 QY 496 SGFERLVGALA-----SALCQLPLRVAVDLWSRELSSAGPVAMFHAQR-----OT 542  
 Db 1038 RRADELLAVLAEEPSALCR-----AVERWARDERPARHRAAVTGLRTAPHARSADRT 1092  
 QY 543 LOEGGVVLLFSP-----GAVALCSEWLQDGVSGPAGHGP-----DAFRA----- 583  
 Db 1093 LIRHAALLVLAGSDSPRLGGALAL-----VQDPCRDRHLPAALDLFAACDPYLP 1145  
 QY 584 -SLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPALFR-----TVPVFTLPSQLPDFLGAL 638  
 Db 1146 SAVAAALPTH-----PEPVLEAFRALLGPDAGEALRLADATTALT-HRVAALVGR 1198  
 QY 639 QOPRAPSGRLOERAEQVSRAL--OPALDSYFHPGT-----PAPRGVGPAGPGAGD 690  
 Db 1199 VTERPETAGHL---AAYVDRRLDRDPAPRAVLLPLVTRLLDDGPEPARAA--LAGVLAAD 1253  
 QY 691 G 691  
 Db 1254 G 1254  
 RESULT 8  
 Q96M86 PRELIMINARY; PRT: 1021 AA.  
 ID Q96M86  
 AC Q96M86  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CNA FLJ32752 FIS, CLONE TEST12001661, WEAKLY SIMILAR TO DYNEIN BETA  
 DE CHAIN, FLAGELLAR OUTER ARM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

TISSUE-TESTIS:  
 RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,  
 RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,  
 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,  
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,  
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,  
 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
 RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,  
 RT "NEDO human cDNA sequencing project."  
 RL Submitt (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK057314; BAB71423.1;  
 SQ SEQUENCE 1021 AA; 112942 MW; A1207BD978EA4D0D CRC64;  
 Query Match 3.1%; Score 116; DB 4; Length 1021;  
 Best Local Similarity 24.0%; Pred. No. 0.72;  
 Matches 103; Conservative 75; Mismatches 276; Indels 228; Gaps 44;  
 QY 55 PAPGPVLAPTHLOTELVLRCQKETDCDCLRVAVHLAVHGWEEPEDEEFGGAANDSGVE 114  
 Db 294 PAPGPPEPLSLKQLILWRVLRPE---CLAGALADFTTSLGRLPLDENTY---APTMPF 347  
 QY 115 EPRNASLQAOVLSFQAYPTARC-VLLEVOVPAALVQFGSGSVVYDCEAALGSE--- 170  
 Db 348 KHSQATQPMILLPPPGHPSATLHPLTVIOKLAAYOQOQKQV---IIIIII 401  
 QY 171 ---VRIWSYTPRYEKELNHTQQLPALPWLNVSDGDNVHLVNSEEQHFGLSLYNQV 227  
 Db 402 PVSVVYSTLSQAYE---GH-----WLVL-----DNCHLM----- 428  
 QY 228 QGPKPRWHKNTLGPQITLNTDLPCLCIQVWPLEDPSVRTNICFREDPRAHONLW- 286  
 Db 429 ---PHWPKELQLLELGRKAVVADLESEQLLDQPE---RNVSTVHRDFR---LWL 477  
 QY 287 ---QAARLRLTLQOS---W-----LLDAPCSLPAEALCRAAGDPGCPQPLV 327  
 Db 478 IVPASASLPAVLTOHSMVPVFNQSLGELGHVLID---SVELAOOVLYHQP---PTOAL- 530  
 QY 328 PPLSWENVTVDKVLEPPLKGHENLQVQVNSSEKILQOECLWADSL----- 373  
 Db 531 PLLLLHGLLHRLQYTRLOAHGRWNSQVTLTVQVLOTQDLWA-SLSNPRAMQELAASV 589  
 QY 374 ---GPLKD---DVLLLETRG-----PODNRSLCA-LEP-----SGCTS 404  
 Db 590 FYGGLGDTEDREALISLTQACLSPPSGSVWQPHTPQSLLATLMLPELRELDAMAECKA 649  
 QY 405 ---LPSKASTRAARLGE-----YLLQDQSGOCLOLWDDDLGALWACPMKY 448  
 Db 650 QMHLLSPPEPRLCGSEGPQAWLLRRQSRALLSALQSRSPVWVPESSRRGAQLA---ERR 706  
 QY 449 IHKRWLVWLACLLFAAALSLLKDKHAKA-----AARGAALLYSADDSGFER 500  
 Db 707 LROR--LVQVNRRLLESLODLTHVIRODESDAPWSVLGPNARRPLEGLVLETEALE--LSQ 762  
 QY 501 LVGALASAL-CQL-----PLR-----VAVDLWSRR-----ELSAQGPVAMFHA 537  
 Db 763 LVGTLQRDLDCLLQLKGAPPCPSRRCAVAHALWTGRLPLPNRPHAPAGQPWHWL-- 820  
 QY 538 QRRQTLOEGGVVLLFSPGAVALCSEWLQDGVSGPAGHGP-DAFRASLSCVLPDFLOGR 596  
 Db 821 --RQLSRRGOLLVRYLGVA-----DASSDVPVRFHLSAFRHRPRLLLA--LRGE 867  
 QY 597 A-----PGSVYVACFDRL---LHPDAVPALFRIV--PVFTLPSQ--LPDFLGA 637  
 Db 868 AALDQNVPSNFFGSR--GSVSSQLQYKRLMNSNPLFRVENGPNPTVPERGULLIGLV 926  
 QY 638 LQOPRAPSGRLOERAEQVSRALQPALDSYFHPGT---PAP 676  
 Db 927 LHAENPDIAGALQDSSQSPSPPLPPVIS-TQAPGTSDLFAP 967  
 RESULT 9

```
Q9EYF6
ID Q9EYF6 PRELIMINARY; PRT; 745 AA.
AC Q9EYF6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 82.6 KDA PROTEIN.
GN YEHQ.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7;
RX MEDLINE=20564182; PubMed=11111050;
RA Yokoyama K., Makino K., Kubota Y., Watanabe M., Kimura S.,
RA Yutsudo C.H., Kurokawa K., Ishii K., Hattori M., Abe H., Iida T.,
RA Yamamoto K., Hayashi T., Yasunaga T., Honda T., Sakakawa C.,
RA Shinagawa H.;
RT "Complete nucleotide sequence of the defective Sakai-VT1 prophage
RT coli O157:H7 derived from the Sakai outbreak.";
RL Gene 258:127-139(2000).
DR EMBL; AP000400; BAB19549.1;
KW Hypothetical protein.
SQ SEQUENCE 745 AA; 82619 MW; 37EC82409B811F43 CRC64;

Query Match 3.1%; Score 114.5; DB 2; Length 745;
Best Local Similarity 21.6%; Pred. No. 0.63;
Matches 163; Conservative 69; Mismatches 236; Indels 287; Gaps 43;

QY 21 LERLVGPDAH-----CSPGLSCLWSDILCLPGLDIPVAPGPPVLATHLQTELVLRC 74
Db 110 LEELATLPDARKRAQVLVAKGITIELF-----CAPGEIPSAKLP-MSDVRFYSRSIRF 163
QY 75 QKETDC-----DLCLRVAVHLAVHGHWEPEDEKPF-----GGAADSGVEEPRNA 119
Db 164 AR-CDICIEGLCEHV--LAVQAFVAKTQQAFFTHLQMRSEHVTSSDDPFASEGNA 220
QY 120 SLQ-----AQVLSFQA-----YPTARCVLLEVVQVPAALVQFGQSGSVV 159
Db 221 CROYVQQLSQALWLGIGISQPLIHYEAAFSRAQQAERCNWR--WVSESLRQLRASV-- 274
QY 160 YDCFEAA-----LGSEVRWSTQRYEKELNHTQQL-----PALPWLNVSDG-- 203
Db 275 -DAFHARASHYHAGECLRLQALNSR-----LNCQAEARRDSVGEVPPVFWRTVVGSGIA 329
QY 204 -----DNVHLV-LNV---SEQHGLSLYWNQVQGPKPWRHKNLTGPQIITLNLHTDLVP 254
Db 330 GEAKLDHLRLVSLGMRQWDIEHYGLRIWF-----LNCQAEARRDSVGEVPPVFWRTVVGSGIA 371
QY 255 CLCIQVWPLEPDSVRTNICPFREDPAHONLWQAAARLRLTLTQSWLL-----DAPCS 306
Db 372 -----RSWRSEQENSPAATRRFLFSQAGALAGGQIVSQAAKRS 410
QY 307 LPAAALCWAPGDPCCQPLVP--PLSWENVTVDKVLEPPL-----LKGHPNLC 353
Db 411 ADGELLATR-----NRLSSVPLSPDAW-----RMLSAPLRQPGIVALEYLRQRPSC 460
QY 354 V-----QVNSSEKQLQEQCL-----WADSLGPKDLDVLLLETRGPDNRSLCALPESGCTSLP 406
Db 461 IRPLNQVDNLFLTPVAECISLQWDSRQTL--DAQVISGEQ-EDN--LLTL-----SLP 509
QY 407 SKASTRAA--RIGEVLLQDLSGQCL-----OLWDDD----- 436
Db 510 ASAPYAVERMAA--LLQOTDPPVCLVSGFVSFVQGLTLEPQVMNMTKTRAWALDAETAP 568
QY 437 ----LGALWACPMKYIHKRWALWLAACLLFAAALSLLILLKKD--HAKAAAGRAALLL 490
Db 569 VVASLPSASVLPVPSTAHO-----LLMRCQALLTQLLHNGRWRYOQSAIGQAEILIA 619
```

```
QY 491 YSADDSGFERLVGALASALCOLPLRVAVDLMSRRELSAQGPVAFHAFQROTLOEGGVVV 550
Db 620 NDLTAVGFYRLAHVL-----GOFRTNESEARVEA 648
QY 551 LLFSPGVALCSEWL-----QDVGSPGAHGDPAFRASLSCVLDP-----FLQGR 596
Db 649 M-----NNGVLLCEQLFPLLOQOGLNRPGE--PGEFCICELR--LPDHRFRWKNHKLFL--- 698
QY 597 APGSYVGACEDRLHLHPDAVPALFRTVPVFTLPSQL 631
Db 699 -----LLPEEVPAPFAIDVCYTSPTTL 721

RESULT 10
Q9HZE5
ID Q9HZE5 PRELIMINARY; PRT; 1193 AA.
AC Q9HZE5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA3063.
GN PA3063.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Pham K., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004730; AG06451.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1193 AA; 135097 MW; 1FE9553F34D3C191 CRC64;

Query Match 3.1%; Score 114.5; DB 16; Length 1193;
Best Local Similarity 21.7%; Pred. No. 1.2;
Matches 154; Conservative 53; Mismatches 240; Indels 263; Gaps 38;

QY 81 DLCLRVAVHLAVHGHWE-----EPEDBEKFGGAADSGV---EPRNASLQAVV--- 126
Db 293 DLRLNMAAG-RVERAWEVGQELLALRPEDRTLADLARLGWTGNGPRAIGFWKQLLAGA 351
QY 127 -----LSQAYPTARCVLLEVVQVPAALVQFGQSGSVVYDCEAALGSEVRW- 174
Db 352 DDPALREHAWRLSLQMFDFDSATELLAPIGAQRQMTDELDALVYSHETRGTPPEGEAWL 411
QY 175 -SYTOPYEKE-----LNHTQQLPALP--WLVNSADGDNVHLVNLVNSEQHGLS 221
Db 412 RGTVQ-RYPQRANORLQOILEHTQOQLEETGVWMAR-----HPFLSVKERMQWA-E 464
QY 222 LYNQVQGPKPWRHKNLTGPQIITLNLHTDLVPCLCIQVWPLEPD-----SV 268
Db 465 THNLFED--PQAW-KVLGVDTPAIRPEFWRLRAALAWALEODDDARAAYERMLALDI 521
QY 269 RTN-----ICPFRED--PRAHONL---WQAAH-----LRLTLQSWLLDAPC--SLPA 309
Db 522 RLNSRDEDLIALYRDSNPKNQALQVLIGSMQSRDRPRLASALQALENLHDWPALKSLLA 581
QY 310 EAALCWAPGDPCCQPLVPPLSW-----ENVTVDKLEPFLKKGHPNLC 353
Db 582 EAEGLPDAQS-----PYVWARLARLAEQEGHGDVAERLYREALVRFP----- 624
QY 354 VQVNSSEKQLQEQCLW-----ADSLGP-----LKDGV-----LLETRGP 388
```

[illegible]

```

RESULT 11
Q96DW2
ID Q96DW2 PRELIMINARY; PRT; 657 AA.
AC Q96DW2:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIMILAR TO RECQ PROTEIN-LIKE 4 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH, AND LYMPHOMA;
RA Strausberg R.;
EL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013277; AAH13277.1; -.
SQ NON_TER 1
FT SEQUENCE 657 AA; 73415 MW; CF51B910F560CA18 CRG64;

```

Query Match	3.0%;	Score 112.5;	DB 4;	Length 657;
Best Local Similarity	19.6%;	Pred. No. 0.78;		
Matches 115;	Conservative 73;	Mismatches 194;	Indels 205;	Gaps
Qy 16	PWLSLERLVGPQDATHCSPGLSC---	RLWDSDL--CLFGDIVPAPGPVLAPTHLQ---	67	
	: : : :         :       :	:       :       :		
Db 68	PCYLRVCKVLRRMGVHCFLGLTATARTTASDVAQHLAAVEEPDLGGPAPVPTNLHLVS	127		
Qy 68	-----TELVLRCQKETCDCLCLRVAVHL--AVHGHWEEPEDEE	103		
	: : : : :   :   : :   :   :   :	:   :   :   :   :   :		
Db 128	SMDRDTQALLTLQGRFQNLDSIIYCNRRDTE--RIAALLRCTLHAAWPVS---	181		
Qy 104	KFGGAOSGVVEEPRNASL-----QAQVLSFQAY-----	PTARCVLLEV	142	
	:   :   :   :   :   :   :	:   :   :   :   :   :		
Db 182	--GGRAPKTTAEAYHAGMSCRRRRVQAFMQLRVVVATVAFGMLDRDPVRAV--LHL	238		
Qy 143	QVPAALVQFGOSGVSVYDFEAAALGSSEVRWSYTOPRYE--KELNHTQOLPALPWLNV	200		
	: : :   :     :   :   :   :   :	:   :   :   :   :   :		
Db 239	GLPPSFSEYVQAVGRAGRD-----GQPAHCHLFLOQGEDRLRRHHVHADSTDFLAVK	292		
Qy 201	ADGDNVHLVLNV-----SEQHFGLSLYNNQVQGP--PKPRW-----	HKN	238	
	:   :   :   :   :   :   :	:   :   :   :   :   :		
Db 293	-----RLVQRVFPACTCTCTRPPEQE-----GAVGGERPVPKYPPQEAQLSHQA	338		
Qy 239	LTGPOIITLNTFDLVPCLCIQVWPLEPDSVFRNICPFREDPRAHONLWQAERLL--T	295		
	: :   :   :   :   :   :   :   :	:   :   :   :   :   :		
Db 339	APGPRRVCMGHERALPTOLTVOALDMPEAETILLCYL-----ELHPHHW-----	LELLATT	390	

[illegible]

```

RESULT 12
Q96F55 PRELIMINARY; PRT; 744 AA.
AC Q96F55;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE SIMILAR TO REQD PROTEIN-LIKE 4 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011602; AAHL1602.1; -.
FT NON_TER 1
SQ SEQUENCE 744 AA; 82839 MW; 6AE0F90340A2C9F1 CRC64;

```

Query Match	3.0%;	Score 112.5;	DB 4;	Length 744;		
Best Local Similarity	19.6%;	Pred. No. 0.92;				
Matches 115;	Conservative 73;	Mismatches 194;	Indels 205;	Gaps		
QY	16	PWLSLERLVGPQDATHCSGPLSC---	RLWDSDL-	CLPGDIVPAPCPVLAPLTHQ---	67	
DB	155	PCYLRCVKVLBERMGVHCFGLTATATRTASDVAQH	LVAEEPD	LHGPAPVPTNHL	SV 214	
QY	68	-----	TELVLRCQKETCDICLRVAVHL--	AVIGHWEEPEDEE	1030	
DB	215	SMDRDTQALLTLQGRKFNLDIIYCNRRDETE---	RIAA	LLRCLHAA	WPGS--- 2468	
QY	104	KFGGAADSGVEEPRNASL-	-----	QAOVLSFOAY-	-----PTARC	VLEV 142
DB	269	---GGRAPKTTAEYHAGCMCSRRRRVQAF	FMQGLRVVVATVAF	CMGLDRDPD	VRV-LHL 325	
QY	143	QVPAALVFGSGSVGVVYDCFEAALGSEVR	TWSYTPRYE--	KELNHTQOLPALP	WLNVS 2000	
DB	326	GLPPSFESYQAVGRAGRD-----	GPAHCHLFLPQGED	RLRELRHVHAD	STDFLAVK 379	
QY	201	ADGDNVHLVNW-	-----	SEEQHFGLSYWN	VOGP-KPRW-----HKN 238	
DB	380	-----	RLVRVFPACTCTCTRP	SEQE-----	GAVGGERIVPKY	PQDEAQLSHQA 425
QY	239	LTGPQIITLNHTDLVPC-LCIVQWPLE	SDSVRTNICPFRED	PRAHQLNQAAR	LRLL--T 295	
DB	426	APGPRRCVMGHERALPIQLTQALDMP	EATETLLCYL---	ELHPHHW---	LELLATT 477	
QY	236	LQSWLLDAPCSLPAAALCWRA	P-----	GGDP	CQ-----PLVPPLSHENV 335	
DB	478	YTHCRINCPGPPAQIQLAHLA	RCPLPVLAVLAQ	LPEDPQGS	SSVEFDMKVLWDSMGWELA 537	
QY	336	TVDKVL-----	EEPLLKGHPNL	CVGVNSSEK	LOLOEC	LWADSLG 374

```
Db 538 SVRALCQLQMDHEPTGVRGVLVEFSELAFLHRLSPGDLTAEEKDQCIDFLY-----G 593
QY 375 PLKDDVLLLETRGPQDNRSKALE-----PSGCTSLPSKASTRAAR-----LGEY 419
Db 594 RVQ-----ARERQALARLRRFQAFHSVAFFPSGCGLEQOQDEERSTRLLKDLGRY 643
QY 420 L-----LQDLQS---GOC-LQLWDDDLGALWACPMKDYIHKR 452
Db 644 FEEEGEGPGMEDAQPEPGQARLQDWEQV-----RCDIRQFSLR 686

RESULT 13
QYRY89
AC Q9RY89 PRELIMINARY; PRT; 478 AA.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHETICAL 51.2 KDA PROTEIN.
GN DR0061.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN STRAIN=R1;
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001869; AAF09658.1;
DR TIGR; DR0061;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 478 AA; 51231 MW; 5E74CDEB8FBA5C95 CRC64;

Query Match 3.0%; Score 111.5; DB 16; Length 478;
Best Local Similarity 25.7%; Pred. No. 0.62;
Matches 105; Conservative 39; Mismatches 145; Indels 119; Gaps 20;

QY 276 REDPRAHQLWQAARLLTLQSWLL-----DAPCSLPAAEALCWRAPGGDPQPLV 327
Db 110 RADPREPLLEAAQ---RIRWLARLARAPQLPGTVPAHSVLY---AQEASALS 161
QY 328 PPLSWENVTVKVEPLPKGHPNLCVQVNSSEKLOQECWADSLGPKDKDDVLLLETRG 387
Db 162 APAAGDGLT-----LHLGGPD---RLAHWQGLRLS---WR-SLCPNNQLLVQDETQS 206
QY 388 PDNRSKALEPSSGCTSLPSKASTRAARLGEYLQDLQSCQCLQDLWDDDLGALWACPMK 447
Db 207 PVSGGQALLRP-----DLFASERQELFTVGGQQLQVLFSG-----DY 244
QY 418 YIHKRWLWLAACLLFAAALSILLLKDKHAKAAARGA-ALLYSADDSGFERLVGALA 506
Db 245 VLLRRRA-----DAAQTRHL-----ARLAALGRACALLLLPAEQHRLRLATLA 289
QY 507 SALCQLPLRVAVDLNRSRRELSAQGPVAFHQAQRQTLQEGGVVLLFSPGAV----- 558
Db 230 RRLRGDPPR-ADDPASGQVAFATPAECCLAAARR-SLQRLEPLLGRFSPAQAHAQAQV 347
QY 559 -----ALCSEWLQDGVSGP---GAHCPHDAFRASLSCVL---PDF 592
Db 348 ALLHLPPHYAAVQQAEEHAASAAEPLEAATDPLVREGAGPPVLPATGHFVIFPLGPD 407
QY 593 LQGRAGSVYGCDFRL--LHPDAVPALFRTVPVFTLPSQLPDFLGL 638
```

```
Db 408 LEVRLPG-----DRLLTLRPD-----YRAELVAVLPQGAATVGDGL 443

RESULT 14
Q66031
AC Q66031 PRELIMINARY; PRT; 1279 AA.
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE IMMEDIATE EARLY GENE PRODUCT (HYPOTHETICAL 136.8 KDA
DE PROTEIN).
GN IRI.
OS Cercopithecine herpesvirus 9 (simian varicella virus), and
OS Cercopithecine herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=35246, 35245;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Cercopithecine herpesvirus 9 (simian varicella virus);
RX MEDLINE=96434997; PubMed=8837883;
RA Gray W.L., Gusick N.J., Ek-Kommonen C., Kempson S.E., Fletcher T.M.;
RA "The inverted repeat regions of the simian varicella virus and
RT varicella-zoster virus genomes have a similar genetic organization.";
RL Virus Res. 39:181-193(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Cercopithecine herpesvirus 7;
RX Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
RA "Complete Sequence of the Simian Varicella Virus Genome.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U33499; AAB39965.1;
DR EMBL; AF275348; AAG27252.1;
DR EMBL; AF275348; AAG27249.1;
DR InterPro; IPR000923; Copper_blue.
DR PROSITE; PS00196; COPPER_BLUE; UNKNOWNL.
SQ SEQUENCE 1279 AA; 136763 MW; A3691FC354F9C877 CRC64;

Query Match 3.0%; Score 111; DB 12; Length 1279;
Best Local Similarity 21.0%; Pred. No. 2.5;
Matches 119; Conservative 39; Mismatches 174; Indels 234; Gaps 28;

QY 266 DSVRTNCPFPREDPRAHQLWQAARLLTLQSWLLDAPCSLPAAEA-LCWRAPGGDP- 323
Db 713 DGPSTNGC-FRRVP-----AGALHTP--VPSEARRAYCAPNVIESL 751
QY 324 --QPLVPLSWENVTVKVEPLPKGHPNLCVQVN-----SSEKLQLEPC 367
Db 752 VDEPMFPP-AWR-----PVLRFDPAAVAEISARGHGVDRFRFPVPSVDALRRRC 800
QY 368 LWADSLGPKDKDDVLLLETRGPQDNRSKALEPSSGCTSLPSKASTRAARLGEYLQDLQSG 427
Db 801 AWMHQVAD-PEDVRLVIYDPLPGESLC-----GPTCEGGAS-----RDVH-- 841
QY 428 QCLQWDDDLGALWACPMKDYIHKRWALWLAACLLFAAALSILLLKDKHAKAA----- 481
Db 842 ----WSDCRGGL-----SVVLAALSNRLCLPOTHAWAGNSGPP 876
QY 482 ----ARGRAALLYSADDSGFERLVGALASALCQLPLR-----VAVDLWSR----- 523
Db 877 DVSQLNANG----ILLSTRDLAFAGAEYLGSRSLASAGRLLILDVAPERPRDGPALS 933
QY 524 -----RELSAQGPVAFHQAQRQTLQEGGVVLLF-----PGAVA----- 559
Db 934 QYHVYVKAAPRPAQAVVRWPR-----TEAGLARAVFASSRTFGPGSFARVETAFENL 987
QY 560 -----LC-----SEWLQDGVSG----- 571
Db 988 YPNEQPIRLCRGANVCYVWATRAGPRTRVPLSPRDYRQVLPFGDFGCKDIARQAQGLAVG 1047
```

QY 572 -----PGAHPGPHDAFRASLSCVL-PDFL-QGRAPGSYVGACFDRLLLHPDAVPAFLFRTVP 623  
Db 1048 AAEVDDAAASHRAANRWGLGAPLPVLPGRPGS-AGP-----RPSDVPLWARVFC 1100  
QY 624 VFTLPSPDLFLGALQOPRAPRSGLRQRAQVSRALQALDYSYFHPGTPA-----675  
Db 1101 RHALPEPDAEPILPPVAGRSVAVFCMSEARESL-PPIPRIILWPPGFGSETAVEMG 1159  
QY 676 -----PGRGVGPGAGGAGD 690  
Db 1160 DGTRLVFEGHGPREDDEGGAAPGCCD 1185  
RESULT 15  
Q9N015  
ID Q9N015 PRELIMINARY; PRT; 757 AA.  
AC Q9N015;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE BRAIN PROTEIN.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=CEREBELLUM;  
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
RA Suzuki Y., Sugano S., Hashimoto K.,  
RT "Isolation of full-length cDNA clones from macaque brain cDNA  
RT libraries."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB046632; BAB03550.1;  
SQ SEQUENCE 757 AA; 78303 MW; 5222A33F169B586B CRC64;

Query Match 3.0%; Score 110; DB 6; Length 757;  
Best Local Similarity 20.9%; Pred. No. 1.5;  
Matches 143; Conservative 53; Mismatches 213; Indels 274; Gaps 33;  
QY 143 QVPAALVQFGSGSVYDCEFAALGSEVRINWSYTPQRYEKELNHTQQLPALPWLNVSD 202  
Db 21 QDPAA-----SDVGSV-----PPVEVYISOEPGAQPDALARSPPAE-----58  
QY 203 GDNVHLVNLVSEEQHFGLSLYWQVQPKPRW-----HKNLTGQ-----243  
Db 59 -----ELPVETPRAGAEVSEVSSSPLPKQVDLPDAKDSGPGPMDPPASEAPDGGPF 112  
QY 244 ---ITLNHTDLVPCLCIQVWPLEPDSVR-----TNICPFREDPRAHONLWQAARL 291  
Db 113 KPEIATMNGAD-----PISQVRGAVEAPGTPKPLIPGSDPGPAAN-----155  
QY 292 RLLTLQSWLLDAPC-SLPAEALCWRAPGDPCQPLVPLSWENVTVDKVLEFPLKGGH 350  
Db 156 -----RTESPMGALQPDAAEW--PGRQSHPPAPP-----184  
QY 351 NLCVQVNSSEKILQECVLWADSLGPKDDVLLLETRGPDQDNRSICALEPSGCTSLPSKAS 410  
Db 185 ---VOAPSTSRGLRVTTRALG-----QPPPEPSASSMASAPA--SSPPANAT 230  
QY 411 TRAARLGE-----YLLQDLSQGCQLWDDDDIGALWACPMCKYIHK 451  
Db 231 APPLRWGPLRRVLSFWSLHVYGVVFLFPALLA-----LAALAAAPAG----P 276  
QY 452 RWALVWLACLFALFAALSILILKKDHAKAAARG-RAALILYSADDSGFERLYGALASALC 510  
Db 277 RLALVAALVLVVASGLRSAYMLTDPYSGOARLGVGRGLVLYNLP---FPLLLTALA-ALT 332  
QY 511 QL-----PLR-----VAVDL---NSRRELSAQG-PVAMFHAORROT 542  
Db 333 LILGAGLPPPLQNPPLLGLGVALVHGVLATDILLSTWNSVLNLLMQGLSCAW-----384

QY 543 LOEGGVVVLLFSPCAVALCSEWLQDGVSG-PGAHGPBDAFRASLSCVLPDFLQ-----594  
Db 385 ---GAAVAL-----GTCLCRRRLDDGPRGWDASPGPRLLAVALGALLASGLQAAALWL 437  
QY 595 -----GRAPGSYVG-----ACFDRLLHPDAV 615  
Db 438 YPGPGRVGRFSWAWGVHFWLRLLELTWALALALAAVAAAPRPPTHEACWAKLMRL-AC 496  
QY 616 PALFRTVPVFTLPSQLPDFLFGALQOQPRAP-----RSGRLQERAPQVSRAL--QPALDSYF 668  
Db 497 PA-----PS-----GKSEVPERPNNCYARPSVSGAGSLDISKSLIRNPAENGQL 540  
QY 669 HPPGTPAPGRGVGPGAGGAGDG 691  
Db 541 ATPSSGAWGSAASLGRGPQGGPG 563

Search completed: May 19, 2002, 15:19:08  
Job time: 6744 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 14:10:19 ; Search time 43.62 seconds  
(without alignments)  
614.258 Million cell updates/sec

Title: US-09-608-918-2

Perfect score: 3716

Sequence: 1 MPVPWFLLSLALGRSPVLS.....TPAPRGVGVGAGPGAGDGT 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137.5	3.7	3707	1	PGBM_MOUSE
2	118.5	3.2	1711	1	PTPO_RAT
3	113.5	3.1	761	1	PQOF_KLEPN
4	112.5	3.0	1208	1	KCQ4_HUMAN
5	108.5	2.9	925	1	W70T_HUMAN
6	107	2.9	825	1	IL4R_HUMAN
7	104.5	2.8	975	1	CUT1_CANFA
8	103.5	2.8	676	1	UL06_HSV1
9	102	2.7	922	1	W70T_MOUSE
10	101.5	2.7	1151	1	ATC7_YEAST
11	100.5	2.7	753	1	UL06_HSVB
12	99	2.7	955	1	CHRD_HUMAN
13	99	2.7	1011	1	M3K6_HUMAN
14	98.5	2.7	587	1	NDC2_RAT
15	98.5	2.7	741	1	TLE4_RAT
16	98.5	2.7	766	1	TLE4_HUMAN
17	98.5	2.7	948	1	CHRD_MOUSE
18	98.5	2.7	1505	1	CUT1_HUMAN
19	98.5	2.7	1524	1	X133_HUMAN
20	98	2.6	1487	1	ICP4_HSVB
21	98	2.6	1487	1	ICP4_HSVBK
22	98	2.6	2205	1	POLN_RUBVT
23	97.5	2.6	418	1	NER3_RAT
24	97.5	2.6	850	1	NRG2_HUMAN
25	97.5	2.6	4393	1	PGBM_HUMAN
26	97	2.6	2594	1	7LES_DROVI
27	96.5	2.6	392	1	GAP_BLVAV
28	95.5	2.6	505	1	CXAA_MOUSE
29	95	2.6	862	1	CUT1_RAT
30	94.5	2.5	614	1	YEHQ_ECOLI
31	94	2.5	993	1	AFSR_STRCO
32	94	2.5	1446	1	IE18_PRRKA
33	93.5	2.5	642	1	FZD1_MOUSE

075325 homo sapien  
Q06418 homo sapien  
P29915 paracoccus  
Q9u449 homo sapien  
P49000 rattus norv  
Q9rca8 bacillus ha  
P41846 caenorhabdi  
Q99558 homo sapien  
Q07954 homo sapien  
Q60755 homo sapien  
Q9up38 homo sapien  
P53564 mus musculu

34 93.5 2.5 713 1 GACL\_HUMAN  
35 93.5 2.5 890 1 TYO3\_HUMAN  
36 93 2.5 672 1 NOO3\_PARDE  
37 92 2.5 428 1 NER3\_HUMAN  
38 92 2.5 553 1 MIS\_RAT  
39 92 2.5 632 1 GIDA\_BACHD  
40 92 2.5 843 1 YO96\_CAEEL  
41 92 2.5 947 1 M3KE\_HUMAN  
42 91.5 2.5 4544 1 LRPL\_HUMAN  
43 91 2.4 368 1 GALT\_HUMAN  
44 91 2.4 648 1 FZD1\_HUMAN  
45 91 2.4 1395 1 CUT1\_MOUSE

#### ALIGNMENTS

RESULT 1  
PGBM\_MOUSE  
ID PGBM\_MOUSE STANDARD; PRT; 3707 AA.  
AC Q05793;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Basement membrane-specific heparan sulfate proteoglycan core  
DE protein precursor (HSPG) (Perlecan) (PLC).  
GN HSPG2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RX MEDLINE=92078153; PubMed=1744087;  
RA Noonan D.M., Fullie A., Valente P., Cai S., Horigan E., Sasaki M.,  
RA Yamada Y., Hassell J.R.;  
RT "The complete sequence of perlecan, a basement membrane heparan  
RT sulfate proteoglycan, reveals extensive similarity with laminin A  
RT chain, low density lipoprotein-receptor, and the neural cell adhesion  
RT molecule.";  
RL J. Biol. Chem. 266:22939-22947(1991).  
RN [2]  
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=89034110; PubMed=2972708;  
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogeli G., Sasaki M.,  
RA Yamada Y., Hassell J.R.;  
RT "Identification of cDNA clones encoding different domains of the  
RT basement membrane heparan sulfate proteoglycan.";  
RL J. Biol. Chem. 263:16379-16387(1988).  
CC -!- FUNCTION: THIS PROTEIN IS AN INTEGRAL COMPONENT OF BASEMENT  
CC MEMBRANES. IT IS RESPONSIBLE FOR THE FIXED NEGATIVE ELECTROSTATIC  
CC CHARGE AND IS INVOLVED IN THE CHARGE-SELECTIVE ULTRAFILTRATION  
CC PROPERTIES. IT INTERACTS WITH OTHER BASEMENT MEMBRANE COMPONENTS  
CC SUCH AS LAMININ AND COLLAGEN TYPE IV AND SERVES AS AN ATTACHMENT  
CC SUBSTRATE FOR CELLS.  
CC -!- SUBUNIT: PURIFIED PERLECAN HAS A STRONG TENDENCY TO AGGREGATE IN  
CC DIMERS OR STELLATE STRUCTURES.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.  
CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED  
CC AND O-LINKED OLIGOSACCHARIDES.  
CC -!- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.  
CC -!- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.  
CC -!- SIMILARITY: CONTAINS 15 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 SEA DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	DR	EMBL; M77174; AAA39911.1; -	FT	DOMAIN	1866	1954	IG-LIKE C2-TYPE DOMAIN 4.
CC	DR	EMBL; J04054; AAA39899.1; -	FT	DOMAIN	1955	2049	IG-LIKE C2-TYPE DOMAIN 5.
CC	DR	EMBL; J04055; AAA39912.1; -	FT	DOMAIN	2050	2148	IG-LIKE C2-TYPE DOMAIN 6.
CC	DR	HSSP; Q07954; 1CR8.	FT	DOMAIN	2149	2244	IG-LIKE C2-TYPE DOMAIN 7.
CC	DR	MGB; MGI:96257; Hsp92.	FT	DOMAIN	2245	2343	IG-LIKE C2-TYPE DOMAIN 8.
DR	DR	InterPro; IPR000561; EGF-like.	FT	DOMAIN	2344	2436	IG-LIKE C2-TYPE DOMAIN 9.
DR	DR	InterPro; IPR000742; EGF_2.	FT	DOMAIN	2437	2532	IG-LIKE C2-TYPE DOMAIN 10.
DR	DR	InterPro; IPR001438; EGF_II.	FT	DOMAIN	2533	2619	IG-LIKE C2-TYPE DOMAIN 11.
DR	DR	InterPro; IPR003006; Ig_MHC.	FT	DOMAIN	2620	2720	IG-LIKE C2-TYPE DOMAIN 12.
DR	DR	InterPro; IPR003598; Ig_C2.	FT	DOMAIN	2721	2809	IG-LIKE C2-TYPE DOMAIN 13.
DR	DR	InterPro; IPR002172; LDL_recept_A.	FT	DOMAIN	2810	2895	IG-LIKE C2-TYPE DOMAIN 14.
DR	DR	InterPro; IPR000034; Laminin_B.	FT	DOMAIN	2896	2980	IG-LIKE C2-TYPE DOMAIN 15.
DR	DR	InterPro; IPR002049; Laminin_EGF.	FT	DOMAIN	2984	3162	LAMININ G-LIKE 1.
DR	DR	InterPro; IPR001791; Laminin_G.	FT	DOMAIN	3163	3241	EGF-LIKE.
DR	DR	InterPro; IPR000082; SEA.	FT	DOMAIN	3245	3425	LAMININ G-LIKE 2.
DR	DR	Pfam; PF00047; Ig; 15.	FT	DOMAIN	3518	3705	LAMININ G-LIKE 3.
DR	DR	Pfam; PF00052; laminin_B; 3.	FT	SITE	65	67	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00053; laminin_EGF; 8.	FT	SITE	71	73	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00054; laminin_G; 3.	FT	SITE	76	78	HEPARAN SULFATE (POTENTIAL).
DR	DR	Pfam; PF00057; ldl_recept_a; 4.	FT	SITE	3615	3617	MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL).
DR	DR	Pfam; PF01390; SEA; 1.	FT	DISULFID	199	212	BY SIMILARITY.
DR	DR	PRINTS; PR00010; EGFBLD.	FT	DISULFID	206	225	BY SIMILARITY.
DR	DR	ProDom; PD003031; Laminin_B; 3.	FT	DISULFID	219	234	BY SIMILARITY.
DR	DR	SMART; SM00180; EGF_Lam; 7.	FT	DISULFID	285	297	BY SIMILARITY.
DR	DR	SMART; SM00001; EGF_Like; 6.	FT	DISULFID	292	310	BY SIMILARITY.
DR	DR	SMART; SM00408; IGC2; 14.	FT	DISULFID	304	319	BY SIMILARITY.
DR	DR	SMART; SM00192; LDLa; 4.	FT	DISULFID	325	337	BY SIMILARITY.
DR	DR	SMART; SM00281; Lamb; 3.	FT	DISULFID	332	350	BY SIMILARITY.
DR	DR	SMART; SM00282; LamC; 3.	FT	DISULFID	344	359	BY SIMILARITY.
DR	DR	SMART; SM00200; SEA; 1.	FT	DISULFID	375	394	BY SIMILARITY.
DR	DR	PROSITE; PS00022; EGF_1; 8.	FT	DISULFID	388	403	BY SIMILARITY.
DR	DR	PROSITE; PS01186; EGF_2; 5.	FT	DISULFID	428	479	BY SIMILARITY.
DR	DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 11.	FT	DISULFID	766	780	BY SIMILARITY.
DR	DR	PROSITE; PS00025; LAM_G_DOMAIN; 3.	FT	DISULFID	792	811	BY SIMILARITY.
DR	DR	PROSITE; PS01209; LDLRA_1; 4.	FT	DISULFID	795	811	BY SIMILARITY.
DR	DR	PROSITE; PS00068; LDLRA_2; 4.	FT	DISULFID	814	829	BY SIMILARITY.
DR	DR	PROSITE; PS00024; SEA; 1.	FT	DISULFID	816	839	BY SIMILARITY.
DR	DR	Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;	FT	DISULFID	842	851	BY SIMILARITY.
DR	DR	Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;	FT	DISULFID	1159	1168	BY SIMILARITY.
DR	DR	Extracellular matrix; EGF-like domain.	FT	DISULFID	1161	1175	BY SIMILARITY.
FT	CHAIN	POTENTIAL.	FT	DISULFID	1178	1187	BY SIMILARITY.
FT	CHAIN	BASEMENT MEMBRANE-SPECIFIC HEPARAN	FT	DISULFID	1190	1206	BY SIMILARITY.
FT	CHAIN	SULFATE PROTEOGLYCAN CORE PROTEIN.	FT	DISULFID	1209	1224	BY SIMILARITY.
FT	CHAIN	SEA.	FT	DISULFID	1211	1234	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 1.	FT	DISULFID	1237	1246	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 2.	FT	DISULFID	1249	1263	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 3.	FT	DISULFID	1275	1287	BY SIMILARITY.
FT	CHAIN	LDL-RECEPTOR CLASS A 4.	FT	DISULFID	1277	1293	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 1.	FT	DISULFID	1295	1304	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 1 (N-TERMINAL).	FT	DISULFID	1307	1322	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 1 (DOMAIN III A).	FT	DISULFID	1563	1572	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 1 (C-TERMINAL).	FT	DISULFID	1565	1579	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 2.	FT	DISULFID	1582	1591	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 3.	FT	DISULFID	1594	1610	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 4 (INCOMPLETE).	FT	DISULFID	1613	1628	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 5 (N-TERMINAL).	FT	DISULFID	1615	1638	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 5 (C-TERMINAL).	FT	DISULFID	1641	1650	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 5 (C-TERMINAL).	FT	DISULFID	1653	1668	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 6.	FT	DISULFID	1792	1839	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 7.	FT	DISULFID	1886	1932	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 8.	FT	DISULFID	1976	2021	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 9 (N-TERMINAL).	FT	DISULFID	2073	2118	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 9 (C-TERMINAL).	FT	DISULFID	2170	2215	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 9 (C-TERMINAL).	FT	DISULFID	2268	2313	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 10.	FT	DISULFID	2365	2413	BY SIMILARITY.
FT	CHAIN	LAMININ EGF-LIKE 11.	FT	DISULFID	2456	2506	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 2.	FT	DISULFID	2554	2599	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 3.	FT	DISULFID	2641	2686	BY SIMILARITY.
FT	CHAIN	IG-LIKE C2-TYPE DOMAIN 3.	FT	DISULFID	2831	2876	BY SIMILARITY.



Query Match 3.7%; Score 137.5; DB 1; Length 3707;  
Best Local Similarity 19.5%; Pred. No. 0.031; ...  
Matches 163; Conservative 73; Mismatches 235; Indels 365;

QY	4	PWFLLSIALGRSPVLSLERLVGQDAT-----HCSF-----	GLSC-----	39		
Db	272	POFLLPSPVPGS-----ACGPQEAASCHGCIIPRDYLCDOGDECRDGDGLGCASPP		323		
QY	40	-----RLWDSIDLCLPGDI-----VPAGGPVLAPTHLOTELV		71		
Db	324	PCPENEFACENGHCALKLWRCD-----GDFCEDERTDEANGCVKQPGVCCGTHFQCQVST		378		
QY	72	LR-----COKETDCDCLCLRAVAVHLAVGHWEPEDEKEKFGGAADSVGEPRNASLQAQ		124		
Db	379	NRCIPASPHCDESDC-----PDRSDEFGCMPPQVVTTPQO-SIGA-		418		
QY	125	VVLSFQAYPTARCVLLEVOVPAAALVQFGQSGVSVYDCFEAALGSEVRIMSYTOPRYEKE		184		
Db	419	-----SRQG---TWFTTCVATGVPTPIIW-----R		441		
QY	185	LNHTTQQLPALPWLNVSADGDNVHLVNLVSEBOHG-----LSLYW		224		
Db	442	LN-WGHIPAHPRVMTSBSGGRTLIIRDVKADQAGAYTCEAMNSGMVFGIPDGVLELV-		499		
QY	225	NOVGPPKPRMKNLTGPOIITLNTHTDLVPCLCIQVMPLEPDSVRTNCP-----FREDPR		280		
Db	500	--PQGPCPCDGHFYLE-----DSASCLPCFCFV-----TNVCQSSLRFRDQIR		541		
QY	281	A---HQNLMQARLRLITLQSWLLDAPCSLPAEALCWRAPGDPQCPQLVPLPSWENVTV		337		
Db	542	LSFDQPNDFKGVNTM-----PSQPGVPPLSSTQLQI		573		
QY	338	DKVL-EPPLLKGHNLGVNVSSEKLQOECLWA-----DSLQ-----		374		
Db	574	DPALQEFQL-----VLSRRFLVHDFAFWALPKQFLGNKVDSYGGFLRYKVRYELA		623		
QY	375	-----PL-KDDVLL-----LETGRQPDNRSICALPESGCTSLPSKASTRAARLGEYLLQ		422		
Db	624	RGMLPEVQKPDVILVGAGYRLHSRGHPH-----PCTLNQRQVOLSE-----		666		
QY	423	DLQSGQCLQDWDDDLGALWACPMKYIHKRWALVWLACLLFA-AALSILILLKKDPAKAA		481		
Db	667	-----EHWVHESGRPVQRAEMLOALASLEAVLLQTQVYNTKMA		703		
QY	482	ARGRAALLYSADSGFERLYGALASAL--COLPLRVAVDLWSRELNSAQGPVAFWFAQR		539		
Db	704	SVGLSDIVM---DTVTHTTTIHRASHVEECRCPIGYS-----GLSCESCAHF-----		749		
QY	540	RTLQEGSVYLLFS-----PGAVALC-----SEWLQDGVSGPGAHGPHDAFR		582		
Db	750	--TRVPGGPLYLTCGCNCNGHASSCDPVYGHCLNCQHTEGPCDCKCKPFFG--DATK		805		
QY	583	ASLSCVLPDFLOGRAPSGYGA-----CFDRLLHPDAVPALFRVTVPVFT---LPSQLP		632		
Db	806	ATATACRP-----CPCPYIDASRRFSDFCF---LDTDQATCDACAPGYTGRRCESCAP		856		
QY	633	DFLGALQOP-----RAPSGRLQRAEQ-----VSRALQPALDSYFH		669		
Db	857	GYEGNPQPGKCRPTTQETIIVRCBERGSLGTSGETCRCKNNVWVGRCLNCECSGDSGFH		912		
RESULT 2						
PTPO_RAT ID PTPO_RAT STANDARD; PRF; 1711 AA.						
AC Q64612;						
DT	01-NOV-1997	(Rel. 35, Created)				
DT	01-NOV-1997	(Rel. 35, Last sequence update)				
DT	15-JUL-1999	(Rel. 38, Last annotation update)				
DE	Osteotesticular protein tyrosine phosphatase precursor (EC 3.1.3.48)					
DE	(Oste-TPP).					
OS	Rattus norvegicus (Rat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

FT DOMAIN 563 652 FIBRONECTIN TYPE-III 7.  
 FT DOMAIN 563 741 FIBRONECTIN TYPE-III 8.  
 FT DOMAIN 742 830 FIBRONECTIN TYPE-III 9.  
 FT DOMAIN 831 921 FIBRONECTIN TYPE-III 10.  
 FT DOMAIN 1150 1418 PROTEIN-TYROSINE PHOSPHATASE 1.  
 FT DOMAIN 1419 1711 PROTEIN-TYROSINE PHOSPHATASE 2.  
 FT ACT\_SITE 1350 1350 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 117 117 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 570 570 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 663 663 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 851 851 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 970 970 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 982 982 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 1711 AA; 187292 MW; 0F04D2D1A47A18A0 CRC64;

Query Match 3.2%; Score 118.5; DB 1; Length 1711;  
 Best Local Similarity 18.9%; Pred. No. 0.37;  
 Matches 151; Conservative 81; Mismatches 272; Indels 293; Gaps 37;

Qy 5 WFLLSALG-RSPVLSRLVGPQDATHCSPLGSLRMLDSDILCLPLGDIVPAPGPVLAP 63  
 Db 322 WASWAGLAGDGYVL---KLSGPMESTSTLGPPECN-----AVFPGP-LPP 364

Qy 64 THLOTELVLURCKETDCLCLRVAVHLAVHGWEEPEDEKFGGADSGVEE---PRNA 119  
 Db 365 GHYTLQL-----KVLAPDYADWVEGSTWLAESA 392

Qy 120 SIQAQVV-----LSFOAYTARCVLLEVQVPAALVQFGOSVGS--VYVDCFEAALGS 169  
 Db 393 ALPREVPGARLWLDGLEASKQGRALLYDDAPGSLGNISVPSGATHYIFCGLVPGAHY 452

Qy 170 EVRIWYTPQRYEKELNHTQQLP--ALPWLNVSDAGDNVHLVNLVSEEHFGLSLYXNQV 227  
 Db 453 RVDIASSTGDISQSTGYTSPPLPQSLVSRSSPSD-----LTIANGPA 497

Qy 228 QGP---PKPRWKNLT-----GPOITLHTDLVP--CLCIVWVP---LEPDSV 268  
 Db 498 PQLEGYKVTWHODGQRSQSGDLVDLGPDTLSLTLSKSLVPGSCYTVSAAWAGNLDSDSQ 557

Qy 269 RTNCP-----FREDPRAONLW-----QARLRLTLQ----- 297  
 Db 558 KIHSTRFPAPPTNLSLGFHAQPAALKASWYHPGGRDAPFLRLYRLPLTLESEKVLPRE 617

Qy 298 ---SML-LDAPCSLPAAALCW-----RAPGDPQCPVLPPLSWENVTVDKVLFP 344  
 Db 618 AQNFSAQLTACCEQVQVLTUNGSSRSANATGWP--PSAPTL--VNVTSADPTQLQ 673

Qy 345 LKKGHPNLCVQVNSSEKQLQCLWADS-----LGPLKDDVLLLETRGPQDNRSICAL 397  
 Db 674 VSWAHV-----PGRSRYQVTLQESTRTATSIINGP-----KEDGTSFLGL 714

Qy 398 EPS-----GCTSLPSKASTRAARLGEY-----LLQDLQSGCQLQWLDDDLGALWACP 444  
 Db 715 TPGTKYKVEISWAGPLTAANVSAWYPLIPNELLVSMQAGSAVV----- 761

Qy 445 MDKIYHKRWALVWLACLLFAALSLILLKDKHAKAAARGRALLIYSADDSGFERLVGA 504  
 Db 762 -----NLAWPSGPGQGGAC-----HAQLSDAGHLS----- 786

Qy 505 LASALCQLPLRVAVDLWSRRELSAQGPVAFWFAHQRRQTLQEG--GVVVLLFSPCAVALCSE 563

Db 787 -----WEQPLKLGQELFMLRDLTPGHTISMVRCRAGPLQASTHLVLVLSVEPGPV----- 836

Qy 564 WLQDVGSGPAGHPDAFRASLSCLVLPD-----FLQGRAPGS-----YVGACFDRL 609  
 Db 837 --EDVLCHP-----EATYALNWTMPAGDDVDCLVVVVERLVPGGTHFVFQVNTSGDAL 888

Qy 610 LHPDAVPALFTVPTVFTLPSQLPDFLQALQPPRAPSRLQRAEQVSRALQPALDSTFPH 669  
 Db 889 LLPNLMPPTSYRLSLTVL-----GRNSRWSRAVSLVCSAEAW-H 928

Qy 670 PPGTPAP-----GRGVG 681  
 Db 929 PPELAEPQVELGTGMG 945

RESULT 3  
 PQOF\_KLEPN STANDARD; PRT; 761 AA.  
 AC P27508;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Coenzyme PQQ synthesis protein F (EC 3.4.99.-).  
 GN PQQF.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxID=573;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 418;  
 RX MEDLINE=9221293; PubMed=1313537;  
 RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;  
 RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon";  
 RL Mol. Gen. Genet. 232:284-294(1992).  
 CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ) BIOSYNTHESIS. IT IS THOUGHT THAT THIS PROTEIN IS A PROTEASE THAT CLEAVES PEPTIDES BOND IN A SMALL PEPTIDE (GENE PQA) WHICH WOULD PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH ARE NECESSARY FOR THE SYNTHESIS OF PQQ.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE INSULINASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; X58778; CAA11584.1; -  
 CC PIR; S20458; S20458.  
 CC MEROPS; M16.006;  
 CC InterPro; IPR001431; Peptidase\_M16.  
 CC Pfam; PF00675; Peptidase\_M16; 1.  
 CC PROSITE; PS00143; INSULINASE; 1.  
 CC PQQ; Hydrolase; Metalloprotease; zinc.  
 CC METAL 49 49 ZINC (BY SIMILARITY).  
 CC ACT\_SITE 52 52 BY SIMILARITY.  
 CC METAL 53 53 ZINC (BY SIMILARITY).  
 CC METAL 130 130 ZINC (BY SIMILARITY).  
 CC SEQUENCE 761 AA; 83616 MW; 187182DBC8E939B5 CRC64;

Query Match 3.1%; Score 113.5; DB 1; Length 761;  
 Best Local Similarity 23.4%; Pred. No. 0.32; Indels 167; Gaps 27;  
 Matches 117; Conservative 50; Mismatches 165;

Qy 282 HQNLQWAAARLRLTLQSWLLDAPCSLPAAALCWAPGG---DPCQPLVPPLSWENVTVD 338  
 Db 182 HRTHYVARRQL-----W-LQGQSLALGELAAARFAAGLAAGEAPPAPPL----- 227

```
QY 339 KYLEPPLKHPNLCVQVNSSEKXLOQELWADSLGPKDQVILLETGPDNRSLSALE 398
Db 228 RUGETALQ-----LAVSS-----OPALWRCPLIALSONVTLLR----- 261
QY 399 PSCTSLSKASTRAARIGBYLLQD-----LQSGQCLQWDDDLGALWACPMKVIHKRWA 454
Db 262 -----EFLDEAPGSLMAGLRQRLAGDVALNW-----LYQDRH 295
QY 455 LWWLACLLFAA-----ALSILLLKKDKAKAARGRAALL----- 489
Db 296 LGWLA-LVFASDRPEEDVROITHWLAQLOOTTPQOQHYQLSRRRQALSPDLQRLORA 354
QY 490 ----LYSADDSGFERLVGAL-----ASALCQ-----LPLRVAVDLMSRRLS 527
Db 355 FGPAPGAPPAGFAFCAALQVAPSVSLACQTSVSPGPVATQGFSLPL-----SRWRRRPES 410
QY 528 AQGPVAFWHAQRQTL-----QEGGVVLLFSPG-----AVALCSEWLQDGVVS----- 570
Db 411 DPALAFAPYQAAGDLVAKCPKAAAPLLHLPSGDPPLRLLRPFFYCSPOAEGLRGEQ 470
QY 571 -----GPGAHPHDAFRAS--LSCVLPDFLQGRAPSGVYGACFDRLRLHPDA--VPA 617
Db 471 LRPLLAALRHAGCHGEHWFDSWQLTLQLE--PGRPEAILQAILQALPVASLTPS 528
QY 618 LFTVPVFTLPOLDFELGALQPR---APRSRLQERAEOVSRLQALPDLSYHPHPTGP 674
Db 529 P-ESIAIRHLMALPERLGTSGHKGWLAALAGGSAEDAQWVARQLS-LITAPVNP-PP 585
QY 675 APG---RGVPGGPGAGD 690
Db 586 APAPCRGVERLVPG-GD 603

RESULT 4
RCQ4.HUMAN
ID RCQ4.HUMAN STANDARD; PRT; 1208 AA.
AC O94761;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ATP-dependent DNA helicase Q4 (RecQ protein-like 4).
GN RECQ4 OR RECQ4.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99097344; PubMed=9878247;
RA Kitao S., Ohsugi I., Ichikawa K., Goto M., Furuichi Y., Shimamoto A.;
RT "Cloning of two new human helicase genes of the RecQ family:
RT biological significance of multiple species in higher eukaryotes.";
RL Genomics 54:443-452(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20021764; PubMed=10552928;
RA Kitao S., Lindor N.M., Shiratori M., Furuichi Y., Shimamoto A.;
RT "Rothmund-thomson syndrome responsible gene, RECQ4: genomic structure
RT and products.";
RL Genomics 61:268-276(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
CC -!- DISEASE: DEFECTS IN RECQ4 ARE A CAUSE OF ROTHMUND-THOMSON
CC SYNDROME (RTS). A DISEASE CHARACTERIZED BY DERMATOLOGICAL FEATURES
CC SUCH AS ATROPHY, PIGMENTATION, AND TELANGIECTASIA AND FREQUENTLY
CC ACCOMPANIED BY JUVENILE CATARACT, SADDLE NOSE, CONGENITAL BONE
CC DEFECTS, DISTURBANCES OF HAIR GROWTH, AND HYPOGONADISM.
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. RECQ SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
Cc -----
DR EMBL; AB006532; BAA74453.1; -
DR EMBL; AB026546; BAA86899.1; -
DR MIM; 603780; -
DR MIM; 268400; -
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
KW Hydrolase; Helicase; ATP-binding; Nuclear protein.
FT NP_BIND 502 509 ATP (POTENTIAL).
FT SITE 605 608 DEAD BOX.
SQ SEQUENCE 1208 AA; 133076 MW; CB809A7765AB48A1 CRC64;

Query Match 3.0%; Score 112.5; DB 1; Length 1208;
Best Local Similarity 19.6%; Pred. No. 0.7; Indels 205; Gaps 31;
Matches 115; Conservative 73; Mismatches 73;

QY 16 PVLVSLRVLGPDQTHCPSGLSC---RLWDSIL--CLPGDITVPAPGPVLAPTHLQ--- 67
Db 619 PCYLRVCKVLRRMGVHCFGLGTATATRTASDAQLAVAAEEDPLHGPVPTNLHLSV 678
QY 68 -----TELVLCOKETDCDCLRVAVHL--AVGHWEPEDEE 103
Db 679 SMDRDTQALLTLLOGKRFQNLDSIIYCNRRDTE---RIALLRTCLHAAPVGS--- 732
QY 104 KFGAADSQVEEPRNASL-----QAQVLSFQAY-----PTARCVLLEV 142
Db 733 --GGRAPKTTAEYHAGMCSRRRRVQRAFPMQGLRVVATVAFGMGLORPDVRAY-LHL 789
QY 143 QVPAALVQFGQSVGVYDCFAALGSEVRINSYTPRYE--KELNHTQQLPALPWLNV 200
Db 790 GLPPSFESVQAVGRAGRD-----GQPAHCHFLFQPGEDLRLRRHVHADSTFLAVK 843
QY 201 AGDNNVHLNV-----SEEQHFGLSYLNVOGVP-PKPRW-----HKN 238
Db 844 -----RLQVRVFPACTCTCTRPPEQE-----GAVGGERPVKYPQEAQQLSHQA 889
QY 239 LTGPQIITLHHTDLVPC-LCIQWPLEPDSVTRNICPFREDPRAHONLQAARLRL--T 295
Db 890 APGPRRVCMGHERALPIQLTVOALDMPPEAETLLCYL---ELHPHHW---LELLATT 941
QY 296 LQSWLIDAPCSLPAEALCWAP-----GGDPQC-----PLVPLSWENV 335
Db 942 YTHCLNCPGGAQALQALAHRCPPPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELA 1001
QY 336 TVDKVL-----EPLLKGHNLCVQVNSSEKXLOQELWADSLG 374
Db 1002 SVRRALCQLQWDHEPRTGVRRGTGVLVESELAFLHRSGLDTAEKQDQICDFLY---G 1057
QY 375 PLKDDVLLILETRGPQNRSLSALE-----PSGCTSLPSKASTRAAR-----LGEY 419
Db 1058 RVQ-----AREQALRLRRTFOAFHSVAFFSCGPGCLEQDDEERSTRLLGLRY 1107
QY 420 L-----LQDLQS---GQC-LQLWDDDLGALWACPMKVIHKR 452
Db 1108 FEEEGQEPGGMGAOGPEPQARLQDWDQV-----RCDIRQFLSLR 1150

RESULT 5
W70T.HUMAN
ID W70T.HUMAN STANDARD; PRT; 925 AA.
AC P57737;
DT 16-OCT-2001 (Rel. 40, Created)
```

```
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE 70 kDa WD-repeat tumor rejection antigen homolog.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK025674; BAB15211.1; .
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PF00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS00082; WD_REPEATS_2; 4.
DR PROSITE; PS0294; WD_REPEATS_REGION; 2.
KW Repeat; WD repeat.
FT REPEAT 75 115 WD 1.
FT REPEAT 124 163 WD 2.
FT REPEAT 166 205 WD 3.
FT REPEAT 209 253 WD 4.
FT REPEAT 242 282 WD 5.
FT REPEAT 292 322 WD 6.
FT REPEAT 335 374 WD 7.
FT REPEAT 418 458 WD 8.
SQ SEQUENCE 925 AA; 100574 MW; A393CE973C94F4A9 CRC64;
Query Match 2.9%; Score 108.5; DB 1; Length 925;
Best Local Similarity 21.7%; Pred. No. 1;
Matches 134; Conservative 53; Mismatches 223; Indels 207; Gaps 31;
QY 10 LALGRSPVVLRLERLVGPDQATHCSPLGLSCLWDSIDLCPGDIVPAPGPVLAPTHLQTE 69
DB 93 LATGSADRTVKLWRLPGPGQA-----LPS-----APGVVLCPEDLPVE 130
QY 70 LVLRCQKETDCDLCLURVAVHLAVHGHBEPEDEKEFGAADSVEEPRNASLQAQVVLFS 129
DB 131 -VLOFHPTSDGLVSAAGTVKV-----W-----DAAKQOP----- 160
QY 130 QAYPTARCVLLEQVPAALVQF-----GQSVGSVVYDCFEAALGSEVRWS-YTPQRYE 182
DB 161 -----LFEAAGDLVDSAVNSRDGALVGTACKD-----KQLRIFDPTKPRAS 204
QY 183 K--ELNHTQQLPALPWLNVSDGDNVHLVNVSEBQHFGLSLYWNQVQGPVKPRHKNLT 240
DB 205 QSTQAHENSRDRLAWM-----GTWEHLV-----STGFNQMRREVKLWDTRFF 248
QY 241 GPQITLNTHTDLVPL-----CQWVPLEPDSVSRTNICFPRE- 277
DB 249 SSALASLTLDTSGLGCLPLLPDPSGLLVLAGKGERQLYCYEVVPOQPALSPVTQCULESV 308
QY 278 -----DPAHQNLWQARLRLLTL-QSWLLDAPCSLPAEALCWR--APGDPQCPQPLVP 328
DB 309 LRGAALVPRQALAVMGCEVLRQLSDTAIVPIGVHVPKRAFEHDLFPDPTAGCVPATD 368
```

```
QY 329 PLSW---ENVTVDKVLEFPPLKKGHPNLCVQVNSSEKLOQECINADSLGPLKDDV--LLL 383
DB 369 PHGWAGDNOQVQKVSINPACRPHFS-----FTSCL-VPPAEPLPDTAQPAVM 415
QY 384 ETR-GPDNRSLCALEPSCGTSLSKASTAARL-----GEYLLQDLQSGCQLQWDDDLG 438
DB 416 ETPVGADADASEGFSFSSSLTS-FSTPSSLGPSLSSSTGIGTSLSRLQSL-----LG 468
QY 439 ALWACPMDKYTHKRWALVWLACLFLFAAALSILLIKKDHAKAAAGRAALLLYSADDSGF 498
DB 469 -----PSSKFRHAQGT-----VLHRDSHITNLKG--LNLTTGESDGF 504
QY 499 ERLVGALASALCQLPLRVAVDLWSRRELSSAOGPVAHQAQR-----ROTLQEGGVV 550
DB 505 -----CANKLRVAVPL-----LSSGGQVAVLELRKPRGLPDTALPTLQNGAAT 548
QY 551 LL-----FSPGVALCSE 563
DB 549 DLAWDPDFPHRLAVAGE 565
RESULT 6
IL4R_HUMAN STANDARD; PRT; 825 AA.
ID IL4R_HUMAN AC P24394;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Interleukin-4 receptor alpha chain precursor (IL-4R-alpha) (CD124
DE antigen).
DE GN IL4R OR IL4RA OR 582J2.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=90171849; PubMed=2307934;
RA Idzerda R.L., March C.J., Mosley B., Lyman S.D., Bos T.V.,
RA Gimble S.D., Din W.S., Grabstein K.H., Widmer M.B., Park L.S.,
RA Cosman D., Beckmann M.P.;
RA "Human interleukin 4 receptor confers biological responsiveness and
RA defines a novel receptor superfamily.";
RL J. Exp. Med. 171:861-873(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91120547; PubMed=2278997;
RA Galizzi J.-P., Zuber C.E., Harada N., Gorman D.H., Djossou O.,
RA Kastelein R., Banchereau J., Howard M., Miyajima A.;
RA "Molecular cloning of a cDNA encoding the human interleukin 4
RA receptor.";
RT Int. Immunol. 2:669-675(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RA "Genome duplications and other features in 12 Mb of DNA sequence from
RA human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [4]
RX VARIANTS VAL-75; ALA-400; ARG-431; LEU-436 AND PRO-786.
RP MEDLINE=97224413; PubMed=9070874;
RA Deichmann K., Bardutzky J., Forster J., Heinzmann A., Kuehr J.;
RA "Common polymorphisms in the coding part of the IL4-receptor gene.";
RL Biochem. Biophys. Res. Commun. 231:696-697(1997).
RN [5]
RX VARIANT ATOPIC ARG-576.
RP MEDLINE=98041803; PubMed=9392697;
RA Hershey G.K.K., Friedrich M.F., Esswein L.A., Thomas M.L.,
```

RA Chatila T.A.;  
 RT "The association of asthma with a gain-of-function mutation in the  
 alpha subunit of the interleukin-4 receptor.";  
 RL New Engl. J. Med. 337:1720-1725(1997).  
 RN [6]  
 RP VARIANT ATOPIC ASTHMA VAL-75.  
 RX MEDLINE=98282087; PubMed=9620765;  
 RA Mitsuhashi H., Izuhara K., Mao X.-Q., Gao P.S., Arinobu Y., Enomoto T.,  
 Kawal M., Sasaki S., Dake Y., Hamasaki N., Shirakawa T., Hopkin J.M.;  
 RA "Ile50Val variant of IL4R alpha upregulates IgE synthesis and  
 RT associates with atopic asthma.";  
 RL Nat. Genet. 19:119-120(1998).  
 RN [7]  
 RP VARIANT ATOPIC ASTHMA VAL-75.  
 RX MEDLINE=99322293; PubMed=10390422;  
 RA Noguchi E., Shibasaki M., Arinami T., Takeda K., Yokouchi Y.,  
 Kobayashi K., Imoto N., Nakahara S., Matsui A., Hamaguchi H.;  
 RA "No association between atopy/asthma and the Ile50Val polymorphism of  
 RT IL-4 receptor.";  
 RL Am. J. Respir. Crit. Care Med. 160:342-345(1999).  
 RN [8]  
 RP VARIANTS PRO-503 AND ARG-576.  
 RX MEDLINE=99250314; PubMed=10233717;  
 RA Kruse S., Japha T., Tedner M., Sparholt S.H., Forster J., Kuehr J.,  
 Ra Deichmann K.A.;  
 RA "The polymorphisms S503P and Q576R in the interleukin-4 receptor alpha  
 RT gene are associated with atopy and influence the signal  
 transduction.";  
 RL Immunology 96:365-371(1999).  
 RN [9]  
 RP VARIANT ALA-752.  
 RX MEDLINE=20143377; PubMed=10677312;  
 RA Ober C., Leavitt S.A., Tsalenko A., Howard T.D., Hoki D.M., Daniel R.,  
 Newman D.L., Wu X., Parry R., Lester L.A., Solway J., Blumenthal M.,  
 RA King R.A., Xu J., Meyers D.A., Bleeker E.R., Cox N.J.;  
 RA "Variation in the interleukin 4-receptor alpha gene confers  
 RT susceptibility to asthma and atopy in ethnically diverse  
 populations.";  
 RL Am. J. Hum. Genet. 66:517-526(2000).  
 RN [10]  
 RP VARIANT ATOPIC ARG-576.  
 RX MEDLINE=20269830; PubMed=10809862;  
 RA Oiso N., Fukui K., Ishii M.;  
 RA "Interleukin 4 receptor alpha chain polymorphism Gln551Arg is  
 RT associated with adult atopic dermatitis in Japan.";  
 RL Br. J. Dermatol. 142:1003-1006(2000).  
 RN [11]  
 RP VARIANT PRO-786.  
 RX MEDLINE=21405389; PubMed=11513543;  
 RA Andrews R.P., Burrell L., Rosa-Rosa L., Cunningham C.M.,  
 RA Brzezinski J.L., Bernstein J.A., Khurana Hershey G.K.;  
 RA "Analysis of the Ser786Pro interleukin-4 receptor alpha allelic  
 RT variant in allergic and nonallergic asthma and its functional  
 consequences.";  
 RL Clin. Immunol. 100:298-304(2001).  
 RN [12]  
 RP VARIANT ILE-579.  
 RX MEDLINE=21182061; PubMed=11285129;  
 RA Lozano F., Places L., Vila J.M., Padilla O., Arman M., Gimferrer I.,  
 RA Suarez B., Lopez de la Iglesia A., Miserachs N., Vives J.;  
 RA "Identification of a novel single-nucleotide polymorphism (Val554Ile)  
 RT and definition of eight common alleles for human IL4RA exon 11.";  
 RL Tissue Antigens 57:216-220(2001).  
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR INTERLEUKIN-4. A SOLUBLE FORM  
 OF THE IL-4 RECEPTOR MAY REPRESENT A REGULATORY MOLECULE SPECIFIC  
 FOR IL-4-DEPENDENT IMMUNE RESPONSES.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A COMMON GAMMA CHAIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DISEASE: DEFECTS IN IL4R ARE A CAUSE OF ATOPIC ASTHMA.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD124 entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd124.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; X52425; CAA36672.1; -  
 DR EMBL; AC004525; AAC23495.1; -  
 DR PIR; A60386; A60386.  
 DR PIR; A47603; A47603.  
 DR MIM; 147781; -  
 DR InterPro; IPR002996; CRIA.  
 DR InterPro; IPR003961; FN\_III.  
 DR InterPro; IPR003531; Hematopo\_receptor\_S\_F1.  
 DR Pfam; PF00041; fn3; 1.  
 DR SMART; SM00060; FN3; 1.  
 DR PROSITE; PS01355; HEMATOPO\_REC\_S\_F1; 1.  
 KW Receptor; transmembrane; Glycoprotein; Signal; Disease mutation;  
 KW Polymorphism.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 825 INTERLEUKIN-4 RECEPTOR ALPHA CHAIN.  
 FT DOMAIN 26 232 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 233 256 POTENTIAL.  
 FT DOMAIN 257 825 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 34 44 BY SIMILARITY.  
 FT CARBOHYD 74 86 BY SIMILARITY.  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 75 75 I -> V (IN ATOPIC ASTHMA).  
 FT VARIANT 400 400 E -> A (IN DBSNP:1805011).  
 FT VARIANT 431 431 C -> R (IN DBSNP:1805012).  
 FT VARIANT 436 436 S -> L (IN DBSNP:1805013).  
 FT VARIANT 503 503 FTid=VAR\_011659.  
 FT VARIANT 503 503 S -> P (IN DBSNP:1805015).  
 FT VARIANT 576 576 Q -> R (IN ATOPIC PATIENTS).  
 FT VARIANT 579 579 V -> I.  
 FT VARIANT 752 752 FTid=VAR\_011661.  
 FT VARIANT 786 786 S -> A (IN DBSNP:1805016).  
 FT VARIANT 786 786 S -> P (IN 1.8% OF THE POPULATION; IN DBSNP:1805014).  
 FT SEQUENCE 825 AA; 89658 MW; 9F886DF5612297F8 CRC64;  
 SQ  
 Query Match 2.9%; Score 107; DB 1; Length 825;  
 Best Local Similarity 19.2%; Pred. No. 1.2;  
 Matches 155; Conservative 66; Mismatches 284; Indels 304; Gaps 36;  
 QY 41 LWSDILILPGDIVP-----APGPVLAPTHQLQELVLRQKREDCDCLRLVAVHIAVH 93  
 Db 103 LWAGQQLLWKGSEKPSKSHVPRAGNLTVHTNVSDTLTLLTWSNPPPDVNYLNHLYAVN 162  
 QY 94 GHWEPEDEKFGGADSGVEEPRNASLQAQVLSFQAYPTARCVLLE--VQVPAALVQF 151  
 Db 163 -INSENDPAD-----FRIY---NVTLEPSLRIAASTLKS 193  
 QY 152 GQSVGSVVYDCFEALGSEVRINSYTOPRYEKELNHQQLPALPWLNVNADGNVHLVNL 211  
 Db 194 GIS-----YRARVANAQCY-----NTTWSSESPSTKWNHSREPPFQHLHLG 236  
 QY 212 VSEQHFGLSL-----YWNQVQGPCKPR-----WHKNLTPQI 244

```

Db 237 VSVSCIVLAVCLLCVSVITKIKKWDQIPNPARSRLVAILIQQAGSOWERSRQEP 296
Qy 245 ITLNH-----TDLVPLCLCIQVWPFDSVRTNICPPREDPRAHQNLQAAARLLTLQSW 299
Db 297 AKCPHWKCLTKLLPCF-----LEHMKRD-----EDPH-----KAA--KEMPFQGS 336
Qy 300 LLDAPCSLPAEALCWRAPGDCPCPLVPLSWENVYDKVLEFPL-----LKGH 349
Db 337 GKSAMCPVEISKTVLW-----PESIVVRCVVELFEAPVECEEEVEEKEG- 382
Qy 350 PNLCOVQVNSSEK-----LQLECLWADSLGPKDDVLLLETRG-----PODNRSLCA 396
Db 383 -SFCASPESSRDDFOEGREGIVARITESLFLDLG-----EENGFCQODMGESCL 432
Qy 397 LEPSGCTG-----LPSKASTRAARLGRYLQDLQSGQCLOLWDDDLGALWACPMCKYI 449
Db 433 LPPSGSTSAHPWDEFPAGPKAPWGK-----EQPLHEPFPASPTQSPN--- 481
Qy 450 HKRWALVWLACLLFAAALSLLKKHAKAAARGAALLYSADDSGFERLVGALASAL 509
Db 482 -----LTC-----TETPLVIAGNPAYRFSFNSLSQSP 508
Qy 510 QCLPLRAVDLWSRRELS-----AQPVAWFHAQRQRIQLEGVVV 550
Db 509 C--PRELGPDPILARHLHEVEPEMPCVPQLSEPTTVQPEPETWETQILRRNLVQHGAAGA 566
Qy 551 LLFSP--CAVALCSWLOQGVG-----GPCAHGPHDAFRASL--SCVLPDFLQ--GRAP 598
Db 567 PVSAPTSYQYEVHVEGQGTQASAVGLGPGCEAGYKAFSSLLASSAVSPEKCGFGASS 626
Qy 599 GSYVGACDRLL-----HPDAVPALFRIVPVT-----LPSQLPDFLGA- 637
Db 627 GEEGYKPPQDLIPGCGDPAPVP-----VPLFTGLDREPPRSPQSSHLPSSEPHLGLE 681
Qy 638 -----LQQRAPRS-----GRQERAEQVSRAALQALD 666
Db 682 PGEKVEDMPKPLPQEQATDPLVDSLGSGIVYSALTCHLCGLKQHGQEDGGQTPVMAS 741
Qy 667 -----YFHPPTGTPAPGRGVGPGAGP 686
Db 742 PCCGCCCGDRSPPTPLRAPDPSGGVP 770

RESULT 7
CUT1_CANFA
ID CUT1_CANFA STANDARD; PRT; 975 AA.
AC P39881;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE CCAAT displacement protein (Homeobox protein Clox) (Clox-1)
DE (Fragment).
GN CUT1 OR CLOX.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart ventricle;
RA MEDLINE=93161940; PubMed=1363085;
RA Andres V., Nadal-Ginard B., Mahdavi V.;
RT "Clox, a mammalian homeobox gene related to Drosophila cut, encodes
RT DNA-binding regulatory proteins differentially expressed during
RT development."
RL Development 116:321-334(1992).
CC -!- FUNCTION: DNA-BINDING REGULATORY PROTEIN, WHERE THE SMALLER
CC PROTEIN SPECIES PROBABLY REPRESENTS THE ACTIVE DNA-BINDING FORM.
CC MAY FUNCTION AS A DEVELOPMENTALLY REGULATED REPRESSOR OF TISSUE-
CC SPECIFIC GENE TRANSCRIPTION BY PREVENTING THE INTERACTION OF
CC TISSUE-SPECIFIC TRANSCRIPTION FACTORS WITH THEIR COGNATE TARGET
CC SEQUENCES. PROBABLY INVOLVED IN CELL-FATE SPECIFICATION IN DIVERSE

```

```

CC DIFFERENTIATION PROGRAMS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms may be produced by
CC alternative splicing.
CC -!- TISSUE SPECIFICITY: A BROAD PATTERN OF EXPRESSION OBSERVED IN
CC TISSUES OF DIVERSE ORIGINS, SUCH AS CARTILAGE, LIVER, BRAIN, LUNG,
CC HEART AND SKELETAL MUSCLE. THERE ARE 2 DISTINCT PROTEIN SPECIES:
CC THE LARGER ONE (230-250 KDA) IS FOUND MAINLY IN ADULT BRAIN, LUNG
CC AND HEART, AND THE SMALLER ONE (180-190 KDA) PREDOMINATES IN EARLY
CC EMBRYONIC TISSUES.
CC -!- DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED DURING DEVELOPMENT.
CC SMALL PROTEIN SPECIES PREDOMINATE IN EARLY EMBRYOS AND ARE
CC UPREGULATED IN COMMITTED MYOBLASTS AND CHONDROCYTES, BUT
CC DOWN-REGULATED UPON TERMINAL DIFFERENTIATION. LARGE SPECIES ARE
CC DETECTED MAINLY IN ADULT TISSUES AND TERMINALLY DIFFERENTIATED
CC CELLS.
CC -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC HETERODIMERIZATION.
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69017; CAA48782.1; -
CC PIR; S33121; S33121.
CC HSSP; P10037; LAU7.
CC TRANSFAC; T01485; -
CC TRANSFAC; T01985; -
CC InterPro; IPR003350; CUT.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF02376; CUT; 3.
CC Pfam; PF00046; homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Transcription regulation; Homeobox; DNA-binding;
CC Developmental protein; Nuclear protein; Repeat; Repressor;
CC Coiled coil; Alternative splicing.
CC NON_TER 1 1
CC DNA_BIND <1 67 CUT 1.
CC DOMAIN 113 169 COILED COIL (POTENTIAL).
CC DNA_BIND 374 461 CUT 2.
CC DNA_BIND 557 644 CUT 3.
CC DNA_BIND 684 743 HOMEBOX.
CC SEQUENCE 975 AA; 105428 MW; 724243B32C5BCFD9 CRC64;
Qy 3 VPWFLLSLALGRSPVLSRLVLPQDATHCS-----PCLSCRLWSDSLCLPGDIVP 55
Db 190 IPSPMSSSVSYPLALSLLK---PPTAPDTSASTLPNPKKESQADGLDLPGAES 246
Qy 56 APGPVLAPTHLQTELVLRCQKCTDCDLCLRVAVHLAVHGHWE-----PE 100
Db 247 AQG-----VLR-----HVKSGLSGVWKDHWSTVQPERKSAAPPE 283
Qy 101 D--EEKFGAADSQ-----V 113
Db 284 DAKSEAGGTTKEGGGGHGPAASSRDHPHRRSTGRNGPALSPTTQSSLSLTGASRS 343
Qy 114 EEPNRLSLQ--VVLSTFQAYPTARCVLLEVVQVPAALVQFGQSVGVYDCEFAALGSEV 171
Db 344 ETPQNSLPSPSPVPMKPAK-----SVPLTPEQYIYMYQEV 383

Query Match 2.8%; Score 104.5; DB 1; Length 975;
Best Local Similarity 20.4%; Pred. No. 2.3;
Matches 174; Conservative 74; Mismatches 244; Indels 363; Gaps 44;

```

```
QY 172 RIWSTQPRYEKELNHTQOLPALPWLNVSDGDNVHLVNLVSEHQHFGCLSLYWNQVGP- 230
Db 384 DTIELTRQVREK-----LAKNG-----ICQIFGEKVLGLS-----QGV 418
QY 231 -----PKPRWHKNTGPQIITLNTDLPVCLCQVW----- 261
Db 419 SDMLSRPKP-WSK-----LTQKREPIRMQLWNLNGELGQGVLPVQGOQGVPLHS 468
QY 262 -----PLEPDSVRTNCPFRE-----DPAHONLWQAARLLTLQSWLLDAPCSLPAPAA 312
Db 469 VTSIQDPLQGGCVSSSTPKTSASCSPAPESMSSSESVSLSLT-----ELVQPCP-PIET- 523
QY 313 LWRAPGDCPCOPLVPPLSWENVTVDKLEPL-LKGHPNLCVQ--VNSSEKL----- 362
Db 524 -----SKDGKPEPSPDPAS-----DSQATPPLSGHSALSIOELVAMSPELDITYG 573
QY 363 QLOECIWLADSLGP-LKDDVLLETRG-----PODNRSLCALEPSGCTSLPSKAST 411
Db 574 RVKEVLTNNLQORLFGETILGLTQGSVDLLSRPKPHKLSLKGREFP----- 622
QY 412 RAARLGEYLLQDLSQGCQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSIL 471
Db 623 -----VRMQLWLN-----PNN-----VEKILM 639
QY 472 LLKKDHAKAARGAALLLYSADDS-----GFERLVG----- 503
Db 640 DMKRMKAYMKRRHS-----SVDSQCPPEPSVGDYSGASPPQHOLKKPRVVLAPEE 695
QY 504 --ALASALCOLPL--RVAVDLWSRRELSSAOGPVAFHQAQROTLOEGVVVLLFSPGAV 558
Db 696 KEALKRAYQOKPYSPKTIETELNLTSTVWFWNYRSIRRE----- 742
QY 559 ALASEWLQDVGSG-PGA-HGPHDAFRASLSCVLPDFLQGRAPGSYVGAC--FDRLHHPDA 614
Db 743 -LFEEIQAGSQGAGARHSP--SARSS-----GAAPSSGDCDGVAAEGPGA 789
QY 615 VPALFRVTPVFTLPSQLPDFLQALQOQPRAPRSGRQERAEQVSRALQPALDSYFHPPTG- 673
Db 790 ADA-EESAPAAAAKSSQ--GGPAEAAVAP-----EEREE-----APRPAEKRRRPRG 835
QY 674 PAPRGVGPAGPGA 688
Db 836 PGPRGGGGGAPGA 850
RESULT 8
UL06_HSV11 STANDARD; PRT; 676 AA.
AC P10190;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virion protein UL6.
GN UL6.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT Herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE OF 1-80 FROM N.A.
RX MEDLINE=88091053; PubMed=2826807;
RA McGeoch D.J., Dairymple M.A., Dolan A., McNab D., Perry L.J.,
RA Taylor P., Challberg M.D.;
RT "Structures of herpes simplex virus type 1 genes required for
```

```
RT replication of virus DNA."
J. Virol. 62:444-453(1988).
CC -!- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC EHv-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10879; BAA01652.1; -
DR EMBL; X14112; CAA32342.1; -
DR EMBL; M19119; AAA45818.2; -
DR PIR; F28133; WMBEX6.
DR InterPro; IPR002660; Herpes_UL6.
DR Pfam; PF01763; Herpes_UL6; 1.
DR ProDom; PD003210; Herpes_UL6; 1.
SQ SEQUENCE 676 AA; 74091 MW; 8CF9D73C6313FAE9 CRC64;

Query Match 2.8%; Score 103.5; DB 1; Length 676;
Best Local Similarity 22.5%; Pred. No. 1.7;
Matches 102; Conservative 46; Mismatches 153; Indels 153; Gaps 24;

QY 295 TLOSLLDAPCSLPAEAAACWRAPGGD-----PCQPLVPL--SMENVTVDK 339
Db 130 TLRITLLDFAHGLVA---CF-APGGSPSSPPKYIDMTCTCLGLVPLILKKRQEGGVTOG 184
QY 340 VLEFPLLKHPNLCVQVNSSEKYLQOECIWLADSLGPLKDDVLLLETRGPQDNRSCLALEP 399
Db 185 LRAF--LKQHP-----LTRQLATVAEA-----AERAGPGFFELALAFD- 220
QY 400 SGCTSLPSKASTRAA-----RLGEYLLQDL---OSGCQLQLWDDDLGALM---- 441
Db 221 -----STRVADYDRVYIYVNHRRGDWLVDRDPTISGQGECLVLP---PLMTGDR 266
QY 442 -----AC-PMDKYIHKRWALVWLACLLFAAALSILLKKDHAKAARGR 485
Db 267 LVFDSPVQRLFPETIVACHSLREHAH-----VCLRLNTASVKVLGKSDSERGVGA 318
QY 486 AALL--LYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELSSAOGPVAFHQAQROT 543
Db 319 ARVNVKVLGEDD---ETKAGSAASRL---VRLIINKGMRH-----VGINDTVRSYL 365
QY 544 QEGGVVVLLFSPGAVALCSEWLQDVGSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVG 603
Db 366 DEAG-----GHL-----IDAPVDG-----TLPGF--GKGSNRSRG 394
QY 604 ACFDRLHHPDAVPALFRTVPVFTLPSQLPDFL---GALQOPRAPRSG---RLQERAEQV 656
Db 395 AGODGGRAPQLRQAFRTAVVNNINGLEGYINNLFCTIERLRETNAGLATQLQERDEL 454
QY 657 SRALQPALDSYFHPPTGPAPRGVGPAGPGA 690
Db 455 RRATAGALERQQRRAADLAESVTGGCGSRPAGAD 488

RESULT 9
W70T_MOUSE
ID W70T_MOUSE STANDARD; PRT; 922 AA.
AC Q9D2V7;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 70 kDa WD-repeat tumor rejection antigen homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

NCBI\_TaxID=10090;  
 [1]  
 QX SEQUENCE FROM N.A.  
 RN STRAIN=C57BL/6J; TISSUE=Kidney;  
 RC MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Adzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 FT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AK018739; BAB31380.1;  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 6.  
 DR PRINTS; PR00320; GPROTEINBRPT.  
 DR ProDom; PD000018; WD40; 2.  
 DR SMART; SM00320; WD40; 6.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; 1.  
 DR PROSITE; PS0082; WD\_REPEATS\_2; 4.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; 2.  
 KW Repeat; WD repeat.  
 FT REPEAT 75 115 WD 1.  
 FT REPEAT 124 163 WD 2.  
 FT REPEAT 166 205 WD 3.  
 FT REPEAT 209 253 WD 4.  
 FT REPEAT 539 581 WD 5.  
 FT REPEAT 589 629 WD 6.  
 FT REPEAT 632 671 WD 7.  
 FT REPEAT 725 765 WD 8.  
 SQ SEQUENCE 922 AA; 100860 MW; C8E4BC1EE9CFE242 CRC64;  
 Query Match 2.78; Score 102; DB 1; Length 922;  
 Best Local Similarity 21.18; Pred. No. 3.4;  
 Matches 146; Conservative 65; Mismatches 227; Indels 254; Gaps 38;  
 QY 130 QAYPTARCVLL-EYQVPAALVQFGSGVGVYDCEAALGSGSVRIWSTY--OPRYEKELN 186  
 DB 112 EALPSVGVVGLPPELPEVILQFHTVDCVLV----STAGTKVTVMDVAKQPLPLEA- 166  
 QY 187 HTQQLPALPWLNVASGDNVHLVL-----NVSEOHFGLSLYW--- 224  
 DB 167 HKDLVQSAVM---SRDGAIVGTACKDKQLRIFDPRTAQASQSTQAHENNRDLRLAWTGI 223  
 QY 225 -----NOVQGPPEPRKHNLTGPOI--ITLNT--DLVPCV----- 256  
 DB 224 QEHLVSTGFNMQRREAKLMDTRLFSSALASVTLTSPGPLIPLDPSGLVLAKGEN 283

QY 257 ---CIQVWPLEPDSVRTNICPFRE-----DPRAHQNLMOAARLRLTLQSWLL----- 301  
 DB 284 QLYCYEVTPOQPALSPVPTQILENVLRGAALVPRALAVMSCEVLQVLQSLDIAPIISH 343  
 QY 302 -----DAPCSLPAEAALCWAPGDPQCPPLVPPLSWENVTVDKLVLEFLL 346  
 DB 344 HVPKRAVEFHDLFPDTAGSMPSADAHMWA--GD-----NOQVKVSLNPAR 389  
 QY 347 KGHPNLVCQVNSSEKLQQLQECLEWAD---SLGPLKDDVLLLE-----TRGPQDNRLSLC 395  
 DB 390 RPHP-----CFTSSLVPTMEPADVQPAEMPRADTDLSEGFSSPSL- 432  
 QY 396 ALEPSGCTSL-PSKASTR--AARIGEYLLDQSQCCQLQWLDDDLGALWACPMKYIHKR 452  
 DB 433 -MSPSTSSLLPSLSSTSGIGTSPQRSQSLGPGSC-----KFRHTQ 474  
 QY 453 WALVWLACLLFAAALSILLKLLKHAKAAAGRAALLYSADDSGFERLVGALASALCOL 512  
 DB 475 GS-----LLHRDSHTNLKG--LNLTTPGESDGF-----CAN 504  
 QY 513 PLRVAVDLWSRRELSSAOPVAVFAQR-----ROTLOEGGVVLL-----FSPGAVAL 560  
 DB 505 RLRAVAVPL-----LSSGGEVAVLELQKPRLPDTPALPLQNGTAVMDLVDPDPHRLAV 559  
 QY 561 CSEWLQDVGSGAGHPHDAFRASLSCVLPDLQ--RAPGSYVGCACFDRL-----LHPDA 614  
 DB 560 AGE-----DA-RIRLWRVPPGGLNVLTPETVLTGHTKIYSLRFHPLA 603  
 QY 615 VPALFR-----TVPVFTPLPSQLPDLGALQOPRAPRSRGLQERAEQV--SRALQP----- 662  
 DB 604 ADVLASSSYDLTVRIWDL-----QTGAERL-KLQGHQDQIFSLAWSPDGKOLA 650  
 QY 663 --ALDSY---FHPPGTAP-GRGVGPGAGPCA 688  
 DB 651 TVCKDGHVRVYEPSSPLPLQEGGPGEGRGA 682  
 RESULT 10  
 ATC7\_YEAST  
 ID ATC7\_YEAST STANDARD; PRT; 1151 AA.  
 AC P40527;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Potential phospholipid-transporting ATPase 4 (EC 3.6.3.1).  
 GN NE01 OR Y1L048W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Barrrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Delvin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS  
 CC OF ATP COUPLED WITH THE TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).  
 CC LEADS TO NEOMYCIN-RESISTANCE WHEN OVEREXPRESSED.  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
 CC (E1-E2 ATPASES). SUBFAMILY IV.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: Z38060; CAA86174.1; -  
DR SGD: S48431; S48431.  
DR SGO: S0001310; NEO1.  
DR InterPro: IPR001757; E1-E2\_ATPase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00122; E1-E2\_ATPase; 1.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 185 205 POTENTIAL.  
FT TRANSMEM 210 230 POTENTIAL.  
FT TRANSMEM 368 388 POTENTIAL.  
FT TRANSMEM 417 437 POTENTIAL.  
FT TRANSMEM 439 459 POTENTIAL.  
FT TRANSMEM 948 968 POTENTIAL.  
FT TRANSMEM 971 991 POTENTIAL.  
FT TRANSMEM 1021 1041 POTENTIAL.  
FT TRANSMEM 1053 1073 POTENTIAL.  
FT TRANSMEM 1079 1099 POTENTIAL.  
FT TRANSMEM 1110 1130 POTENTIAL.  
FT MOD\_RES 503 503 PHOSPHORYLATION (PROBABLE).  
SQ SEQUENCE 1151 AA; 130217 MW; DC7225CC9577DBE6 CRC64;

Query Match 2.7%; Score 101.5; DB 1; Length 1151;  
Best Local Similarity 21.1%; Pred. No. 4.9;  
Matches 64; Conservative 49; Mismatches 112; Indels 79; Gaps 15;  
QY 102 EEKFGGAADSVPEPRNASIAQVVLVSQAYPTARCVLLEVOVPAALVOFGSGSVVVD 161  
DB 749 EDKLQKQVKSIELLNAGIKIW-MLTGDKVETARC-----VSISAKLISRGQYVHTIT-- 801  
QY 162 CFEEALGSEVRIWSTYOTPRYKELNHTQOLPALPLNLSA-----DGDNVHLVNLNVEEQH 217  
DB 802 -----KVTRP--EGAFNQLYLK-----INRNACLLIDGESLGMFLKHVEQEF 842  
QY 218 FGLSLYNNQV---QGPKPRNH-----KNLGPQIITL-----NHTDLVPLCIIQVW---- 261  
DB 843 FDVVVHLPTVIACRCTPOKADVALVIRKMTGKRVCCIGDGDNDYSMIQCADVGVIQCK 902  
QY 262 -----PLEPDSVRNINCFREDPRAHNLW-----QAARLRLTLQSWLLDAPCSLPA 309  
DB 903 EGKQASLAADFSITQFCHLTE-----LLLWHRNYSYKRSAKLAQFVHRGLIIAICQ--A 955  
QY 310 EAALC-----WRAPGDCQPLVPPLSWENVTVDKVLPEPLKLGHPNLCVQVNS 358  
DB 956 VYSCSLFEPALYQGLWLMVGATCYTMAPVFS-----LTLDHDIIESLTKIYPELYKELTE 1012  
QY 359 SEKL 362  
DB 1013 GKSL 1016

RESULT 11  
UL06\_HSVB  
ID UL06\_HSVB STANDARD; PRT; 753 AA.  
AC P28944;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Virion gene 56 protein.  
GN 56.  
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=31520;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92295566; PubMed=1318606;  
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;

RT "The DNA sequence of equine herpesvirus-1.";  
RL Virology 189:304-316(1992).  
CC !- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA  
CC PACKAGING.  
CC !- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,  
CC EHV-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: M86664; AAB02491.1; -  
DR PIR: B36801; WZBER8.  
DR InterPro: IPR002660; Herpes\_UL6.  
DR Pfam: PF01763; Herpes\_UL6; 1.  
DR ProDom: PD003210; Herpes\_UL6; 1.  
SQ SEQUENCE 753 AA; 83992 MW; C5E118F78BED203 CRC64;

Query Match 2.7%; Score 100.5; DB 1; Length 753;  
Best Local Similarity 22.9%; Pred. No. 3.4; Indels 97; Gaps 20;  
Matches 99; Conservative 46; Mismatches 190;  
QY 295 TLQSWLLDAPCSLPAAALCWRAPGPGQCPQLVPPPLSWENVTVDKVLFFLLKHPNLCV 354  
DB 156 TLQSWLLDFVRSITA---CFSASEPDPGTASFAYIDW--IACGLIPLQLKRAPGATV 209  
QY 355 QVNSSEKLQECLEADSLGDKPLKDDVLLLETRGPQDNRLS-----CALE 398  
DB 210 HPK-----LWR-----KLPTDVPSELES--CVDERDLAGKLYVANSLLREGLEAVVE 253  
QY 399 PSCT-SLPSKASTRAARLGEYLLQDL-----QSGCCLQWD---DDLGLWACPM 446  
DB 254 LARCTASVIMDYDRVNFIFHYTRREVAIDSTGKRGECVLWQPIWKQSGVLFDSPLQ 313  
QY 447 KY-----IHKRWALVWLACLLFAAALSILLKDKHAKAAARGAALLLYSADDSGER 500  
DB 314 RICEVCNCHALREHAKLCOLLNTVPVKILVGRKKDEAQP--GWASKAVDKLWGEEL 371  
QY 501 LVGALASALCOLPLRVAVDLWSRRELSAQGPVAFWHAQRROTLOEGGVVLLSPGVAL 560  
DB 372 HSSSAASRL-----VKLIWNKSMRHIGDITETVRSYLNSTNLISGAQVDTSLP---- 423  
QY 561 CSEWLQDGVSGCAHGP-HDAFRASLSCVLPDFLOGRAGSYVCACFDRLHDPADPALF 619  
DB 424 ---FGSGKTKQGNMVPQEAFTSVINGMLEG-----YVN-----NLF 462  
QY 620 RTVPVFTLPSQLPDFLCAQQAPRSGRLQERAEQVSRALQPALDSYFHPGTPA--PG 677  
DB 463 KTIEDLRTGNS-----GLLDQLR-DRESEITHLREQLLRVSQAADGSGTQPGASSAALPG 516  
QY 678 RGVGPGAGPGAG 689  
DB 517 SGAKSGAG-GLG 527

RESULT 12  
CHRD\_HUMAN  
ID CHRD\_HUMAN STANDARD; PRT; 955 AA.  
AC Q9H2X0; Q9P022; Q9P023; Q9P024; Q9P025; Q9H2W8; Q9H2W9; Q9H2D3;  
AC Q95254;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chordin precursor.  
GN CHRD  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



```
RESULT 13
M3K6_HUMAN STANDARD; PRT; 1011 AA.
ID M3K6_HUMAN
AC O95382;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 6 (EC 2.7.1.1.-).
GN MAP3K6 OR MAPKKK6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99092374; PubMed=9875215;
RA Wang X.S., Diener K., Tan T.-H., Yao Z.;
RT "MAPKKK6, a novel mitogen-activated protein kinase kinase, that
RT associates with MAPKKK3";
RL Biochem. Biophys. Res. Commun. 253:33-37(1998).
CC -!- FUNCTION: ACTIVATES THE JNK, BUT NOT ERK OR P38 KINASE PATHWAYS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF1003118; AAD05304.1;
DR MIM; 604468;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 371 629 PROTEIN KINASE.
FT NP_BIND 377 384 ATP (BY SIMILARITY).
FT BINDING 400 400 ATP (BY SIMILARITY).
FT ACT_SITE 494 494 BY SIMILARITY.
FT ACT_SITE 494 494 BY SIMILARITY.
SQ SEQUENCE 1011 AA; 112512 MW; F7294BA0587D2EC9 CRC64;

Query Match 2.7%; Score 99; DB 1; Length 1011;
Best Local Similarity 20.4%; Pred No. 6,6;
Matches 146; Conservative 93; Mismatches 258; Indels 220; Gaps 39;

QY 113 VEEPRNASLQAQVLSFOAYPTARCVLVQVPAALVQFGSGVYVDFEALGSEVR 172
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 31 VAEQHNVCVHTFALNRRNRPGDRAKALSVLLP--LVQLEGSVAPDLY-CMCGRIYKDF 87
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 173 IWSYTOPRYEKLNTQQLPALPWLNVSDGD-NVH-----LVLVNVEOHF----- 218
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 88 FSSGQ-----DAGHQE--AYHWYRKAFDVEPSLHSGINAVALLLAAGQHFEDSKELRL 140
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 219 -GLSL-----YNNQVQGPKPWRHKNLTGPOIITLNTHTDLVPC-----L 256
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 141 IGKLGCLLARKGCVEKQWYMDV-----GEYLGAQILANDPTQVLLAAEQLYKL 190
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 257 CIQVWPL-----EPDSVTRNICPFREDP-----RAHONLW-----QAARL 291
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 191 NAPIWLVSVNETFLYQHFRTPPEPGGPPRAH--FWLHFLQSQCFKTKACAGDQC 248
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 292 RLLTLQSWLLDAPCSLPAEALCWRAPGDCQP-----LVP-----PLSWENVYDKVL 341
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 249 LVLVLEMNKVLPAKLEVR-----GTDPSVSTVTLSELEPETQDIPSSW----- 291
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 342 EFPLLKGHPNLCVOVNSEKQLQOECLWADSLGPKDDVLLLETRGPQDNRSLCALPSPG 401
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 292 TFPV-----ASIC-GVSASKR--DERCCFLVALPPAQDVQLCFPSVG--HCQWFCGLIQAW 342
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 402 CT-----SLPSKASTRAARLGEYLLQDLSQGCQLQDWDLLCALWACPMKYIHKRWALVW 457
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 343 VTPNDSTAPAEAEAGAGEMLEFDEYETGERLVLGKGTGVVYA-GRDHRTRVRIAKE 401
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 458 LACULFAAALSJLLKKD-----HAKAAARCR-----AALLYSADDSGEERLV--- 502
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 402 IP-----ERDSRFQPLHEETALHRLRHKNIVRYLGSASQCGYLIKIFEE 447
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 503 ---GALASALCOL--PLR---VAVDLWSRRRELSAQGVAMFHAQRROTLOEGGVVLLFS 554
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 448 VPGSLSSLLRSVWGPKLDNESTISFYTRQLQGLVGLHDHNVHRD-IGKDNVLINTFS 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 555 P-----GAVALCSEWLQGV-----SGPGAGHPDAFRASLSCLVLPDF 592
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 507 GLLKISDFGTSKRLAGITPCTETFTGTLOYMAPEIIDQGPGRYG-KAADIWLSLGTCTVIEM 565
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 593 LQGRAP---GSYVGACFD---RLHHPDAVPALFRTVPVETPLPSQLPD----- 633
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 566 ATRGPPEHLEGSPAAMFQVGMVKVHPMPSSLSAEQAFLRTFEPDPLRASAQTLIG 625
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 634 --FLGALQPPAPRSRGLQER-AEQVSRALQPALDSYF-----HPPGTP 674
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 626 DPFLQPKGRSRSPSPRHAPRPSDAPSPTPSANSTQSTQTFPCPOAPSOHPSPSP 682
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
NDC2_RAT STANDARD; PRT; 587 AA.
ID NDC2_RAT
AC P70545;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intestinal sodium/dicarboxylate cotransporter (Na(+)/dicarboxylate
DE cotransporter).
DE Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WISTAR;
RX MEDLINE=97107437; PubMed=8950177;
RA Khatri I.A., Kovacs S.V.B., Forstner J.F.;
RT "Cloning of the cDNA for a rat intestinal Na+/dicarboxylate
RT cotransporter reveals partial sequence homology with a rat intestinal
RT mucin.";
RL Blochim. Biophys. Acta 1309:58-62(1996).
CC -!- FUNCTION: COTRANSPORT OF SODIUM AND DICARBOXYLATES SUCH AS
CC SUCCINATE AND CITRATE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE NADC/P/PRO87 FAMILY OF TRANSPORTERS.
CC NADC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51153; AAB97095.1;
DR InterPro; IPR001898; Na_sulph_symp.
DR Pfam; PF00939; Na_sulph_symp; 3.
DR PROSITE; PS01271; NA_SULFATE; 1.
KW Transport; Transmembrane; Sodium transport; Symport.
```

```

FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
SQ SEQUENCE 587 AA; 63531 MW; F9B74F921BDC8712 CRC64;

Query Match 2.7%; Score 98.5; DB 1; Length 587;
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 88; Conservative 44; Mismatches 163; Indels 145; Gaps 19;

QY 312 ALCWRAPGDCQPLVPPLSWENVVDKLEPPLKGHNLCVQVNS---SEKLQOECL 368
DB 2 ATCW-----PALWAYRFYLIIVLCPLIFLLPLPLIVQTKRAYCAYSIIIMALL 48
QY 369 WADSLGLKDDVL---LLETGPODNRSLCALEPSCCTSLPSKASTRAARLGEYLLQDL 424
DB 49 WCTEALPLAVTVLFPVIVFLLMGIMD-----ASEGLHRLVLSRTPT-----YCLSGA 94
QY 425 QSGCQLQWLDDDLGALWACPMKDIHK-RWALVWLACLLFAAALSILIL----- 471
DB 95 DGGHCRTLEP---AQTHCPPSASYHRSAAALLLGLFMLVTAFLSWISWTATAMVPT 151
QY 472 -----LLKDKHAKAA----- 481
DB 152 GHAVLEQLQASKKQVEGNNNTFELQECPOKEVTKLDNGQPVSAPEPTKTQOEHR 211
QY 482 -ARGRAALLIYSADSGFERLIGALASALCOLPLRVAVDLWSRRELSAOGPVAFHQR 540
DB 212 FSQGLSLCICYSASGGIDNLGTTPNLVLOGVN---SLFKPMQCGELCFVYVWFLPHH 268
QY 541 QTLQGGVVVLFSPGAVALC---SEWL-----ODGVSPGGAHGPDAF 581
DB 269 DHLAAG---LAMATGPIPLGCOLPELWLWGRGTGKACFFGQHPDQVQAAGA---HEFC 322
QY 582 RASLSCVLPDFLQ-----GRAPSGYVG---ACF-----DRLLHPDAVPALFRTVPVFLPSQ 630
DB 323 REDLFTVLVLVLVLFWFTREPFGFFGQGTGFANEGQSMPSDGTVAIFISLWFIIPSK 382
QY 631 LPDFLQALQOPAPRSGRLQ 650
DB 383 IP---GLMEDPKKP--GKLIK 397

RESULT 15
TLEA_RAT
ID TLEA_RAT STANDARD; PRT; 741 AA.
AC Q07141;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Transducin-like enhancer protein 4 (ESP2 protein).
GN TLE4 OR ESP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=94064640; PubMed=8245004;
RA Schmidt C.J., Sladek T.E.;
RT "A rat homolog of the Drosophila enhancer of split (groucho) locus
RT lacking WD-40 repeats.";
RL J. Biol. Chem. 268:25681-25686(1993).
CC -!- FUNCTION: NUCLEAR EFFECTOR MOLECULE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

```

```

CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14463; AAC37640.1;
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 5.
CC PRINTS; PR00320; GPROTEINRPT.
CC SMART; SM00320; WD40; 6.
CC PROSITE; PS00678; WD_REPEATS_1; 2.
CC PROSITE; PS50082; WD_REPEATS_2; 2.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Nuclear protein; Phosphorylation; Repeat; WD repeat.
FT DOMAIN 1 106 GLN-RICH (OR Q DOMAIN).
FT DOMAIN 107 173 GLY/PRO-RICH (OR GP DOMAIN).
FT DOMAIN 175 242 CCN DOMAIN.
FT REPEAT 441 472 WD 1.
FT REPEAT 499 529 WD 2.
FT REPEAT 543 573 WD 3.
FT REPEAT 585 615 WD 4.
FT REPEAT 667 697 WD 5.
FT REPEAT 708 738 WD 6.
FT MOD_RES 218 218 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 233 233 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SQ SEQUENCE 741 AA; 80490 MW; 43F105E568FD87A5 CRC64;

Query Match 2.7%; Score 98.5; DB 1; Length 741;
Best Local Similarity 20.0%; Pred. No. 4.8;
Matches 101; Conservative 59; Mismatches 153; Indels 193; Gaps 26;

QY 23 RLVPQDATHCSGSLSCRLWDSILCLPGDIVPAG---PVLAP---THLQTELVLRCQKET 78
DB 297 RTDAPTGSNSTPLR-----PVPKPPGVDPLASSLRTMAVPCPYPT 340
QY 79 DCDLCLVAVHLAVGHWEPEDEKEFGGAADSGVEEPRNASLQAAQVVLSPQAYPTARCV 138
DB 341 P-----FGIVHAGMNGELTSP-----GAAYAGL---HNISPMQSAAAA----- 376
QY 139 LLEVQVPAALVQFGQSVGVVYDCFEAALGSEVRISWYTPQRYEKELNHTQOLPALP--- 195
DB 377 -----AAAAAAYGRS-----PVGFDPHHMRVPAIPNL 407
QY 196 -----WLVNSADGDNVHLVNLVNSEOHLGLSLVWNVQVGPVPRHKNLTGP--- 242
DB 408 TGIPGKPAYSFHVSADG-----OMQPVPEP---PDALIGPGIP 443
QY 243 -----QIITLHNTDLVPCPL-----CLOVWPLEPDSVVRNICFFREDPRAHQ 283
DB 444 RHARQINTLHNGEVVCAVTSINPRHRYVTGKGVKVKWIDITDPGNKSPVSOL---DCLNRD 501
QY 284 NLWQAARL---RLL-----TLQSWLLDAPCSLPAEAAALCWAPGDCQPLVPPLSW 332
DB 502 NYIRSCRLLPDGRLTVIGGEASTLSIWDLAAP-TPRIKAELTSSAPA---CYALA----- 552
QY 333 ENVTVDKLEFPPLKLGHPNLCVQVNSSEKLOLQECNLWADSLGPKLDDVLLLETGPODNR 392
DB 553 --ISPDSKVCVF-----SCCSGDNIA-----VW-----DLHNQTLVRFQFGHTDGA 590
QY 393 SICALEPSCGCTSLPSKASTRAARLGEYLLQDLQSGCQLQWLWD---DDLGAWACPMKDYIH 450
DB 591 SCIDISNDG-TKLMTGGLDNTR-----SWDLREGLOLQOHDFTSQIFSLGVCPTGE--- 641
QY 451 KRWALVWLACLLFAAALSILILKKD 476

```

Db 642 -----WLAUGMENSNEVLHVTKPD 661

Search completed: May 19, 2002, 17:01:01  
Job time: 10242 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 11:19:39 ; Search time 89.34 seconds  
(without alignments)  
860.343 Million cell updates/sec

Title: US-09-608-918-2  
Perfect score: 3716  
Sequence: 1 MPVPWLLSLALGRSPVLS.....TPAPGRGVGPGAGDGT 692

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3716	100.0	692	22	AA61880 Human cytokine rec
2	3699.5	99.6	705	22	AAU29322 Human PRO polypept
3	3699.5	99.6	705	22	AAU04956 Human Interleukin
4	3699.5	99.6	705	22	AB87606 Human PRO20040..H
5	3699.5	99.6	705	22	AA61884 Chimeric zcytor14
6	3604.5	97.0	675	22	AA61885 Chimeric zcytor14
7	3588	96.6	688	22	AA61883 Chimeric zcytor14
8	2971	80.0	575	22	AA61881 Human variant zcyt
9	1432	38.5	309	21	AA76143 Human secreted pro
10	1305	35.1	267	22	AA88448 Human membrane or
11	1300	35.0	332	22	AA25864 Human protein sequ

12	1079	29.0	204	22	AAE10920 Human gene 12 enco
13	326.5	8.8	617	21	AA76048 Murine skin cell p
14	326.5	8.8	617	22	AB55987 Skin cell protein,
15	309.5	8.3	667	22	AAU04957 Human Interleukin
16	174	4.7	866	17	AAU04185 Human Interleukin
17	174	4.7	866	19	AAW61272 Human IL-17R (hCTL
18	174	4.7	866	20	AAW92409 Human IL-17R prote
19	174	4.7	866	21	AAV99941 Human IL-17R prote
20	174	4.7	866	21	AAV97131 Human Interleukin
21	174	4.7	866	21	AAV97181 Human Interleukin
22	174	4.7	866	21	AA803807 Human Interleukin
23	174	4.7	866	22	AA862066 Human IL-17R (hCTL
24	174	4.7	866	22	AAV72754 Murine Interleukin
25	165	4.4	864	17	AAW04184 Mouse Interleukin
26	165	4.4	864	19	AAW61271 Murine IL-17R prot
27	165	4.4	864	20	AAW92408 Murine IL-17R prot
28	165	4.4	864	21	AAV99935 Murine IL-17R prot
29	165	4.4	864	21	AAV97130 Murine Interleukin
30	165	4.4	864	21	AAV97180 Murine Interleukin
31	165	4.4	864	21	AA803806 Murine IL-17R poly
32	165	4.4	864	22	AA862060 Murine Interleukin
33	165	4.4	864	22	AAV72748 Murine skin cell s
34	162	4.4	330	21	AAV75947 Skin cell protein,
35	162	4.4	330	22	AB55886 A mammalian solubi
36	119.5	3.2	840	21	AAV93652 Soluble Interleuki
37	119.5	3.2	840	21	AAV70028 Novel human diagno
38	118	3.2	1042	22	ABG26706 Osteotesticular pr
39	116.5	3.1	1711	19	AAW70506 Mutant osteotestic
40	116.5	3.1	1711	19	AAW70507 Novel human diagno
41	114.5	3.1	336	22	ABG03482 Novel helicase pro
42	112.5	3.0	1208	20	AAW95050 Human RecQ4 helica
43	112.5	3.0	1208	21	AA820993 Novel human diagno
44	110	3.0	199	22	ABG20476 Mouse oocyte prote
45	110	3.0	664	22	AAE06034

#### ALIGNMENTS

#### RESULT 1

AA61880  
ID AA61880 standard; Protein; 692 AA.  
XX AC AA61880;  
XX DT 08-MAY-2001 (first entry)  
XX DE Human cytokine receptor Zcytor14.  
XX KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
XX KW antiinflammatory; gene therapy; vaccine.  
XX OS Homo sapiens.  
XX PN WO200104304-A1.  
XX PD 18-JAN-2001.  
XX PF 30-JUN-2000; 2000WO-US18383.  
XX PR 07-JUL-1999; 99US-0348854.  
XX (ZYMO ) ZYMOGENETICS INC.  
XX PA Presnell SR, Burkhead SK, Pownder SL;  
XX PI WPI; 2001-112618/12.  
XX DR N-PSDB; AAC85027.  
XX DR New polypeptide encoding a human cytokine receptor Zcytor14, for  
XX PT treating inflammation e.g. rheumatoid arthritis  
XX PS Claim 2; Page 2; 112pp; English.

XX CC The invention provides a new human cytokine receptor designated zcytor14.  
CC zcytor14 can be expressed by standard recombinant methodology. The  
CC encoding nucleic acid is useful for detecting the expression of a  
CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be  
CC used to screen biological samples in vitro for the presence of zcytor14.  
CC proteins, polypeptides and peptides having zcytor14 activity can be  
CC administered to a subject who lacks an adequate amount of this  
CC polypeptide, for treating inflammation and conditions such as rheumatoid  
CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14  
CC antibodies) can be used to treat a subject who produces an excess of  
CC zcytor14. zcytor14 nucleotide sequences can also be used to provide  
CC zcytor14 to a subject. The present sequence represents the human  
CC cytokine receptor zcytor14.  
XX  
SQ Sequence 692 AA;

Query Match 100.0%; Score 3716; DB 22; Length 692;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVPWLLSLALGRSPVLSRLVGPQDATHCSPLGSLRWLSDILCLPGDIVPAGPV 60  
DB 1 MPVPWLLSLALGRSPVLSRLVGPQDATHCSPLGSLRWLSDILCLPGDIVPAGPV 60  
QY 61 LAPTHLQTELVLRCQETDCDLRLVAVHLAVHGHWEPEDEEFKFGAAGSGVEEPRNAS 120  
DB 61 LAPTHLQTELVLRCQETDCDLRLVAVHLAVHGHWEPEDEEFKFGAAGSGVEEPRNAS 120  
QY 121 LQAQVVLSTQAYPTARCVLEQVPAALVQFGSGVSVVYDCFFAALGSEVRIWSYTPR 180  
DB 121 LQAQVVLSTQAYPTARCVLEQVPAALVQFGSGVSVVYDCFFAALGSEVRIWSYTPR 180  
QY 181 YKELNHTQQLPALPWLNLNSADGNVHLVNLVSEQHFGLSLYNQVQGGPKPRWHNLT 240  
DB 181 YKELNHTQQLPALPWLNLNSADGNVHLVNLVSEQHFGLSLYNQVQGGPKPRWHNLT 240  
QY 241 GPQIITLNTDLVPCICVWPLEPSVTRNICPFREDPRAHONLWQAAARLLTLQSWL 300  
DB 241 GPQIITLNTDLVPCICVWPLEPSVTRNICPFREDPRAHONLWQAAARLLTLQSWL 300  
QY 301 LDAPCSLPAEALCWAPGDCQPLVPPLSWENVTVDKVFLLKHPNLCVQVNSSE 360  
DB 301 LDAPCSLPAEALCWAPGDCQPLVPPLSWENVTVDKVFLLKHPNLCVQVNSSE 360  
QY 361 KLQQLBCLWADSLGPKLDDVLLLETRGPQNRSLCALPESGCTSLPSKASTRAARLGEYL 420  
DB 361 KLQQLBCLWADSLGPKLDDVLLLETRGPQNRSLCALPESGCTSLPSKASTRAARLGEYL 420  
QY 421 LODLOSSQCLQWDDDLGALWACPMCKYIHKRWALWLAALLFAAALLSLILLKKHAKA 480  
DB 421 LODLOSSQCLQWDDDLGALWACPMCKYIHKRWALWLAALLFAAALLSLILLKKHAKA 480  
QY 481 AARGAALLYSADDSGFERLVCALASALCOLPLRVAVDLWSRRELSAQGPVAFHQAORR 540  
DB 481 AARGAALLYSADDSGFERLVCALASALCOLPLRVAVDLWSRRELSAQGPVAFHQAORR 540  
QY 541 QTLQEGGVVVLLFSPGAVALCSEWLQDGVSGPAGHPHDAFRASLSCVLPDFLQGRAPGS 600  
DB 541 QTLQEGGVVVLLFSPGAVALCSEWLQDGVSGPAGHPHDAFRASLSCVLPDFLQGRAPGS 600  
QY 601 YGACFDRLHLPDPAVALFTVPVFLPSQLPDFLQALQOPRAPRSRLQERAEQVSRAL 660  
DB 601 YGACFDRLHLPDPAVALFTVPVFLPSQLPDFLQALQOPRAPRSRLQERAEQVSRAL 660  
QY 661 QPALDSYRHPPTPAPGRGVGPGAGPCAGDGT 692  
DB 661 QPALDSYRHPPTPAPGRGVGPGAGPCAGDGT 692

RESULT 2  
AAU29322

ID XX AAU29322 standard; Protein; 705 AA.  
AC XX AAU29322;  
DT XX 18-DEC-2001 (first entry)  
XX XX Human PRO polypeptide sequence #299.  
DE PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200168848-A2.  
XX PD 20-SEP-2001.  
XX PF 28-FEB-2001; 2001WO-US06520.  
XX PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 03-MAR-2000; 2000US-187202P.  
PR 06-MAR-2000; 2000US-186968P.  
PR 14-MAR-2000; 2000US-189320P.  
PR 14-MAR-2000; 2000US-189328P.  
PR 15-MAR-2000; 2000US-189328P.  
PR 21-MAR-2000; 2000US-190828P.  
PR 21-MAR-2000; 2000US-191007P.  
PR 21-MAR-2000; 2000US-191048P.  
PR 21-MAR-2000; 2000US-191314P.  
PR 28-MAR-2000; 2000US-192655P.  
PR 29-MAR-2000; 2000US-193032P.  
PR 29-MAR-2000; 2000US-193053P.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 04-APR-2000; 2000US-194449P.  
PR 04-APR-2000; 2000US-194677P.  
PR 11-APR-2000; 2000US-195975P.  
PR 11-APR-2000; 2000US-196000P.  
PR 11-APR-2000; 2000US-196187P.  
PR 11-APR-2000; 2000US-196690P.  
PR 18-APR-2000; 2000US-196820P.  
PR 18-APR-2000; 2000US-198121P.  
PR 18-APR-2000; 2000US-198585P.  
PR 25-APR-2000; 2000US-199397P.  
PR 25-APR-2000; 2000US-199550P.  
PR 25-APR-2000; 2000US-199654P.  
PR 03-MAY-2000; 2000US-201516P.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 05-JUN-2000; 2000US-209832P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 22-AUG-2000; 2000US-0644848.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
XX XX (GETH ) GENENTECH INC.  
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX WPI: 2001-602746/68.  
XX N-PSDB: AAS46223.  
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
PT presence of tumours, such as prostate and breast tumours, in mammals and  
PT to screen for modulators of the compounds -  
XX

Claim 11; Fig 598; 774pp; English.

Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the animal. Mammals include dogs, cats, cattle, horses, sheep, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

Query Match 99.6%; Score 3699.5; DB 22; Length 705;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPVPFLLSLALGRSPVLSRLERLVGPQDATHCSPGLSCLRLWSDILCLPGDIVPAPGPV 60  
Db 1 mpvpfllslalgrspvlsrlervpqpqathcspglscrlwdsdlclpgdivpapg 60  
QY 61 LAPTHLOTELVLRCOKETDCDCLRVAVLHVHGWEEPEDEKFGGAADSGVVEEPNAs 120  
Db 61 lapthlqtelvircoketdcdclrvavhlavhgwepedeekfggaadsgveepnAs 120  
QY 121 LQAVVLSFQAYPTARCVLLEVVQVPAALVQFGOSGVSVVYDCEFAALGSEVRISWYTPQR 180  
Db 121 lqavvlsfqayptarcvllvqvpaalvqfgsgsvvydcfeaalgsevrIswytqpr 180  
QY 181 YKELNHTQOLPALPMLNVSAGDNVHLVNVSEQHFGLSLVNVOGQPKPRWHKNT 240  
Db 181 ykelnhtqolpalpmlnvsagdnvhlvlnvseeqhfglslyvnvgqppkprwhknt 240  
QY 241 GPQITLNTHTDVLVPCICIOVPLEPDSVRTNICPFREDPAHONLWQAARLRLTLQSWL 300  
Db 241 gpqitlnthtdlvpcicigvplepdsrvrtnicpfredpahonlwqaarlrIrltlqswl 300  
QY 301 LDAPCSLPAEALCWRAPGDCQPLVPPLSWENVTVDKLEPLLKGHNLVCQVNSSE 360  
Db 301 ldapcslpaealCwrapgdcqplvpplswenvtvdKlepllkghnlvcqvnsse 360  
QY 361 KLOLECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYL 420  
Db 361 kIQleclwadsIgpIkddvllletrgpqdnrsIcalepsgctslpskastraarlgEyl 420  
QY 421 LODLQSGCQLQWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSILLKKDHAK 479  
Db 421 lodlqsgcqlqWdddlgalwacpmkdyIhkrwalvwlacllfaaalsIllkkdhakg 480  
QY 480 -----AAARGRAALLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELS 527  
Db 481 wlrlIkdvrsrgaaargraallysaddsgferlvgalasalcolplrvavdlwsrrrels 540  
QY 528 AQGPVAFHQAQRRTLOEGGVVLLFSPGAVALCSEMLQDVGSPGAGHPDFAFRSLSC 587  
Db 541 aggpvafhaqrtrtlqeggvvllfsgavalcsewIqdgvgpgahghpdafrafrslsc 600  
QY 588 VLPDFLQGRAPGSYVACFDRLLHPDAVPALFRTVPVFTLPSQLPDLGALQOPRPSG 647  
Db 601 vlpdfIqgrapgsyvacfdrllhpdavpalfrtvpvftlpsqlpdlgalqopraprg 660  
QY 648 RLQERAEQVSRLQALPDLSYFHPPGTPAPGRGVGPGAGCGAGDGT 692  
Db 661 rIqeraeqvsralqpaldsyfhpptgPAPgrgvpgagpgagdgdt 705

RESULT 3  
AAU04956  
ID AAU04956 standard; Protein; 705 AA.  
XX  
AC AAU04956;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Human Interleukin 17 receptor, IL-17RH2.  
XX  
KW Human; Interleukin-17 receptor; IL-17RH2; agonist; antagonist;  
KW 200040; DNA 164625-2890; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; diabetes mellitus;  
KW allergic disease; asthma; demyelinating disease;  
KW degenerative cartilaginous disorder; transplantation associated disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT /label= Signal\_peptide  
FT Protein 21..705  
FT /label= Mature\_IL\_17RH2  
FT Region 107..112  
FT /note= "N-myristoylation site"  
FT Modified-site 118..121  
FT /note= "Asn is N-glycosylated"  
FT Region 152..157  
FT /note= "N-myristoylation site"  
FT Modified-site 186..189  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 198..201  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 211..214  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 238..241  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 248..251  
FT /note= "Asn is N-glycosylated"  
FT Region 319..324  
FT /note= "N-myristoylation site"  
FT Modified-site 334..337  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 357..360  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 391..394  
FT /note= "Asn is N-glycosylated"  
FT Region 438..443  
FT /note= "N-myristoylation site"  
FT Domain 453..473  
FT /note= "Transmembrane domain"  
FT Region 516..521  
FT /note= "N-myristoylation site"  
FT Region 552..555  
FT /note= "cAMP/GMP-dependent protein kinase phosphorylation site"  
FT Region 583..586  
FT /note= "Glycosaminoglycan attachment site"  
FT Region 612..617  
FT /note= "N-myristoylation site"  
FT Region 692..697  
FT /note= "N-myristoylation site"  
FT Region 696..701  
FT /note= "N-myristoylation site"  
FT Region 700..705  
FT /note= "N-myristoylation site"  
XX  
PN WO200146420-A2.  
XX  
PD 28-JUN-2001.  
XX





DR N-PSDB; AAF92138.  
XX Eighty four nucleic acids encoding PRO polypeptides, useful in  
PT molecular biology, including use as hybridization probes, and in  
PT chromosome and gene mapping.  
XX  
XX  
XX Claim 12; Fig 162; 278pp; English.  
XX  
XX The present sequence is a human PRO polypeptide (secreted and  
CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or  
CC anti-PRO antibodies are useful for preparation of a medication useful in  
CC the treatment of a condition which is responsive to the PRO protein,  
CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be  
CC employed as molecular weight markers for protein electrophoresis. The PRO  
CC coding sequence has applications in molecular biology, including use as  
CC hybridisation probes, and in chromosome and gene mapping.  
XX  
XX  
SQ Sequence 705 AA;  
  
Query Match 99.6%; Score 3699.5; DB 22; Length 705;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
  
Qy 1 MPVPWFLSLALGRSPVLSRLERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPV 60  
Db 1 mpvpwflslalgrspvlsrlervpqqdathcpsglscrlwdsdlclpgdivpapgvp 60  
  
Qy 61 LAPTHLOTELVLRCOKETDCDCLRLVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNAS 120  
Db 61 lapthlqtelvlrcoketdclrlvavhlavhghwepeedeekfggaadsgveeprnas 120  
  
Qy 121 LOAQVWLSFOAYPTARCVLLEVVQVPAALVQFGSGSVVYDCFEALGSEVRWISYQTPR 180  
Db 121 lqaqvwlslfoayptarcvllvvqpaaalvqfgsgsvvydcfeaalgservriwysyqtp 180  
  
Qy 181 YEKELNHTQOLPALPWLNVSDAGDNVHLVNVSEEQHFGLSLYWNOVGPPKPRWHKNTL 240  
Db 181 yekelnhtqolpalpwlrvnsadgdnvhlvlnvseeqhfglsllywnvgppkprwhkntl 240  
  
Qy 241 GPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLAQAARLLTLTQSWL 300  
Db 241 gpqiitlnhtdlvpcliciqvwplepdsvrtnicpfredprahqnlqaarlltltqswl 300  
  
Qy 301 LDAPCSLPAEALCWRAPGDCQPLVPPLSWENVTVDKLEPPLKLGHPNLCVQVNSSE 360  
Db 301 ldapcslpaealcwragpgdcqplvpplswenvtvdklepplkghpnlcvqvnssse 360  
  
Qy 361 KIQLOECLWADSLGPLKDDVLLLETRGPQDNRSICALEPSCGTSLSKASTRAARLGEYL 420  
Db 361 kqlqecclwadslgplkddvllletrgpqdnrsicalepsgtslpskastraarlgelyl 420  
  
Qy 421 LQDLQSGQCQLQWDDDLGALWACPMDBKYIHKRWALVWLACLLFAAALLSLTLKDKDHAK - 479  
Db 421 lqdlsgqcqlqwdldgalwacpmdbyihkrwalvwlacllfaaalsltllkdkdhakg 480  
  
Qy 480 -----AAAGRAALLLYSADDSGFERLVGALASALCOLPLRVAVDLWSRRELS 527  
Db 481 wlrlkqdvrsaaagraalllysaddsgferlvgalasalcqlprvavdlwsrrels 540  
  
Qy 528 ACGPVANFHAQRQTLOEGGVVWVLFSPCAVALCSEWLQDGVSGPGAHGPHDAFRASLSC 587  
Db 541 adgpvawfhaqrqtloeggvvwlfspgavalcsewlgdgvsgpgahgphdfraslsc 600  
  
Qy 588 VLPDFLQGRAGPSYVGCACFDRLHHPDAVPALFRTPVFTLPSOLPDFLQALQOPRAPRS 647  
Db 601 vlpdfqgragsyvgacfdrlhpdavpalfrtpvftlpsolpdfqlaalqoprapsrg 660  
  
Qy 648 RLQERAEQVSRALQALPDLSYFHPGTPAPGRGVGPGAGPGAGGCT 692  
Db 661 rlqeraeqvsralqalpdlsyfhpptpapgrgvpgagpgagdgct 705

RESULT 5  
AAB61884  
ID AAB61884 standard; Protein; 705 AA.  
XX  
AC AAB61884;  
XX  
DT 08-MAY-2001 (first entry)  
XX  
DE Chimeric Zcytor14 protein #2.  
XX  
KW Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
KW antiinflammatory; gene therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN W0200104304-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 30-JUN-2000; 2000WO-US18383.  
XX  
PR 07-JUL-1999; 99US-0348854.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Burkhead SK, Pownder SL;  
XX  
DR WPI; 2001-112618/12.  
XX  
PT New polypeptide encoding a human cytokine receptor Zcytor14, for  
PT treating inflammation e.g. rheumatoid arthritis -  
XX  
PS Claim 2; Page 105-107; 112pp; English.  
XX  
CC The invention provides a new human cytokine receptor designated Zcytor14.  
CC Zcytor14 can be expressed by standard recombinant methodology. The  
CC encoding nucleic acid is useful for detecting the expression of a  
CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
CC used to screen biological samples in vitro for the presence of Zcytor14.  
CC Proteins, polypeptides and peptides having Zcytor14 activity can be  
CC administered to a subject who lacks an adequate amount of this  
CC polypeptide, for treating inflammation and conditions such as rheumatoid  
CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14  
CC antibodies) can be used to treat a subject who produces an excess of  
CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
CC Zcytor14 to a subject. The present sequence represents a chimeric  
CC Zcytor14 protein.  
XX  
SQ Sequence 705 AA;  
  
Query Match 99.6%; Score 3699.5; DB 22; Length 705;  
Best Local Similarity 98.2%; Pred. No. 0;  
Matches 692; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
  
Qy 1 MPVPWFLSLALGRSPVLSRLERLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPV 60  
Db 1 mpvpwflslalgrspvlsrlervpqqdathcpsglscrlwdsdlclpgdivpapgvp 60  
  
Qy 61 LAPTHLOTELVLRCOKETDCDCLRLVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNAS 120  
Db 61 lapthlqtelvlrcoketdclrlvavhlavhghwepeedeekfggaadsgveeprnas 120  
  
Qy 121 LOAQVWLSFOAYPTARCVLLEVVQVPAALVQFGSGSVVYDCFEALGSEVRWISYQTPR 180  
Db 121 lqaqvwlslfoayptarcvllvvqpaaalvqfgsgsvvydcfeaalgservriwysyqtp 180  
  
Qy 181 YEKELNHTQOLPALPWLNVSDAGDNVHLVNVSEEQHFGLSLYWNOVGPPKPRWHKNTL 240  
Db 181 yekelnhtqolpalpwlrvnsadgdnvhlvlnvseeqhfglsllywnvgppkprwhkntl 240  
  
Qy 241 GPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLAQAARLLTLTQSWL 300  
Db 241 gpqiitlnhtdlvpcliciqvwplepdsvrtnicpfredprahqnlqaarlltltqswl 300

Db 241 gpqiitlnhtdlvpcicqvwplesdvrtncifredprahqnlwqaarlrltlqswl 300  
 QY 301 LDAPCSLPAEALCWRAFGDPCQPLVPPLSWENVTVDKVLFFLLKGPMLCQVNSSE 360  
 Db 301 ldapcslpaeaalcwragdpcqplvpplswenvtvdkvlefpplkghpnlcvqnsse 360  
 QY 361 KLQLECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYL 420  
 Db 361 klqlqeclwadslgpklkddvllletrgpqdnrsicalpsgctsipskastaaralgeyl 420  
 QY 421 LODLOSQCLOLWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSLIILLKKDHAK - 479  
 Db 421 lqdlqsgqcilqlddldgalwacpmdkyihkrwalvwlacllfaaalsliillkkdhakg 480  
 QY 480 -----AAAGRAALLYSDDSGFERLVGALASALCQPLRVAVDLMSRRRLS 527  
 Db 481 wlrlkqdvrtsgaaargaallysadsdgsferlvgalasalclprrvavdlmsrrrels 540  
 QY 528 AOGPVAMFHAORRQTOEGGVVLLFSPGVALCSEWLQDGVSGPGAGHPHDAFRASLSC 587  
 Db 541 aggpvawfhaqrqtllqeggvvllfsgpvalcsewlqdgvsppgahghphdafraflsc 600  
 QY 588 VLPDFLQGRAPGVYACFDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQOPRPSRG 647  
 Db 601 vlpdfllqgrapgvvyacfdrlhpdavpalfrtvpvftlpsqlpdlfgalqqpraprg 660  
 QY 648 RLQERAEQVSRALQPALDSVFHPPTAPGRGVGPGAGGAGDGT 692  
 Db 661 rlqeraeqvsralqpaldsyfhpptapgrgvpgagagdggt 705

## RESULT 6

AAB61885  
 ID AAB61885 standard; Protein; 675 AA.

AC AAB61885;

XX 08-MAY-2001 (first entry)

XX Chimeric Zcytor14 protein #3.

XX Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;  
 XX antinflammatory; gene therapy; vaccine.

XX Homo sapiens.

XX WO200104304-A1.

PN 18-JAN-2001.

XX 30-JUN-2000; 2000WO-US18383.

PR 07-JUL-1999; 99US-0348854.

XX (ZYMO ) ZYMOGENETICS INC.

PA Presnell SR, Burkhead SK, Pownder SL;

PI WPI; 2001-112618/12.

DR New polypeptide encoding a human cytokine receptor Zcytor14, for  
 PT treating inflammation e.g. rheumatoid arthritis -

XX Claim 2; Page 107-109; 112pp; English.

XX The invention provides a new human cytokine receptor designated Zcytor14.  
 CC Zcytor14 can be expressed by standard recombinant methodology. The  
 CC encoding nucleic acid is useful for detecting the expression of a  
 CC Zcytor14 gene in a biological sample. Anti-Zcytor14 antibodies can be  
 CC used to screen biological samples in vitro for the presence of Zcytor14.  
 CC proteins, polypeptides and peptides having Zcytor14 activity can be  
 CC administered to a subject who lacks an adequate amount of this  
 CC polypeptide, for treating inflammation and conditions such as rheumatoid

CC arthritis. In contrast, Zcytor14 antagonists (e.g. anti-Zcytor14  
 CC antibodies) can be used to treat a subject who produces an excess of  
 CC Zcytor14. Zcytor14 nucleotide sequences can also be used to provide  
 CC Zcytor14 to a subject. The present sequence represents a chimeric  
 CC Zcytor14 protein.

XX Sequence 675 AA;

Query Match 97.0%; Score 3604.5; DB 22; Length 675;  
 Best Local Similarity 97.5%; Pred. No. 0;  
 Matches 675; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MPVPWFLSLALGRSPVWLSLERLVGPDPDATHCSPGLSCRLWDSITLCLPGDIVPAGPV 60  
 Db 1 mpvfwflslalgrspvwlslerlvgpdpdathcspglscrlwdsitlclpgdivpagpv 60  
 QY 61 LAPTHLOTELVLRCQKEDTCDLCLRVAVHLAVHGWEEPEDEKFGGAADSGVEEPRNAS 120  
 Db 61 lapthlqtelvlrcqketedclclrvavhlavhgwepedeekfggaadsgveeprnas 120  
 QY 121 LQAQVVLVFOAYPTARCVLLEVOVPAALVQFGQSVGVVYDFEALGSEVRWSYTQPR 180  
 Db 121 lqaqvvlvfoayptarcvllvovpaalvqfgqsvgvvydfcealgssevrwsiytqpr 180  
 QY 181 YEKELNHTQOLPALPMLNVSDGDNVHLVNLVSEOHFGLSLYWNQVCGPPKPRWHKNT 240  
 Db 181 yekelnhtqolpalpmlnvsgdgnvhlvnlvseehfglslvwnvqvgppkprwhknt 240  
 QY 241 GPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHONLWQAARLLTLQSWL 300  
 Db 241 gpqiitlnhtdlvpclicqvwplesdvrtncifredprahqnlwqaarlrltlqswl 300  
 QY 301 LDAPCSLPAEALCWRAFGDPCQPLVPPLSWENVTVDKVLFFLLKGPMLCQVNSSE 360  
 Db 301 ldapcslpaeaalcwragdpcqplvpplswenvtvd-----vnsse 343  
 QY 361 KLQLECLWADSLGPKDDVLLLETRGPQDNRSICALPSGCTSLPSKASTRAARLGEYL 420  
 Db 344 klqlqeclwadslgpklkddvllletrgpqdnrsicalpsgctsipskastaaralgeyl 403  
 QY 421 LODLOSQCLOLWDDDLGALWACPMKDYIHKRWALVWLACLLFAAALSLIILLKKDHAKA 480  
 Db 404 lqdlqsgqcilqlddldgalwacpmdkyihkrwalvwlacllfaaalsliillkkdhaka 463  
 QY 481 AARGRAALLYSDDSGFERLVGALASALCQPLRVAVDLMSRRRLSACQPVWAFHQR 540  
 Db 464 aargraallysadsdgsferlvgalasalclprrvavdlmsrrrelsacqpvwafhqr 523  
 QY 541 QTLQEGGVVWLLFSPGVALCSEWLQDGVSGPGAGHPHDAFRASLSCVLPDLQGRAPGS 600  
 Db 524 qtlqeggvvwlflfsgpvalcsewlqdgvsppgahghphdafraflscvlpdlqgrapgs 583  
 QY 601 YVGACFDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQOPRPSRGKLERAEQVSRAL 660  
 Db 584 yvgacfdrlhpdavpalfrtvpvftlpsqlpdlfgalqqpraprgslqeraeqvsral 643  
 QY 661 QPALDSVFHPPTAPGRGVGPGAGGAGDGT 692  
 Db 644 qpaldsvfhpptapgrgvpgagagdggt 675

## RESULT 7

AAB61883  
 ID AAB61883 standard; Protein; 688 AA.

XX AAB61883;

XX 08-MAY-2001 (first entry)

XX Chimeric Zcytor14 protein #1.

XX Cytokine receptor; Zcytor14; human; inflammation; rheumatoid arthritis;

KW antiinflammatory; gene therapy; vaccine.  
XX Homo sapiens.  
XX WO200104304-A1.  
XX 18-JAN-2001.  
XX 30-JUN-2000; 2000WO-US18383.  
XX 07-JUL-1999; 99US-0348854.  
XX (ZYMO ) ZYMOGENETICS INC.  
XX Presnell SR, Burkhead SK, Pownder SL;  
XX WPI; 2001-112618/12.  
XX New polypeptide encoding a human cytokine receptor zcytor14, for  
XX treating inflammation e.g. rheumatoid arthritis -  
XX Claim 2; Page 102-104; 112pp; English.  
XX The invention provides a new human cytokine receptor designated zcytor14.  
XX zcytor14 can be expressed by standard recombinant methodology. The  
XX encoding nucleic acid is useful for detecting the expression of a  
XX zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be  
XX used to screen biological samples in vitro for the presence of zcytor14.  
XX proteins, polypeptides and peptides having zcytor14 activity can be  
XX administered to a subject who lacks an adequate amount of this  
XX polypeptide, for treating inflammation and conditions such as rheumatoid  
XX arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14  
XX antibodies) can be used to treat a subject who produces an excess of  
XX zcytor14. zcytor14 nucleotide sequences can also be used to provide  
XX zcytor14 to a subject. The present sequence represents a chimeric  
XX zcytor14 protein.  
XX Sequence 688 AA;  
SQ

Query Match 96.6%; Score 3588; DB 22; Length 688;  
Best Local Similarity 95.7%; Pred No. 0;  
Matches 675; Conservative 0; Mismatches 0; Indels 30; Gaps 2;

Qy 1 MPVPWFLSLALGRSPVLSRLVLPQDATHCSPGLSCLRLWDSILCLPGDIVPAGPV 60  
Dy 1 mpvpwflslalgrspvlsrlvlpqdaqthcspglscrlwdsilclpgdivpapg 60  
Qy 61 LAPTHLQTELVLRCOKETDCDCLRLVAVHLAVHGHWEPEDEKFGAADSGLVEEPRNAS 120  
Dy 61 laphthlqtelvlrcoketdcdclrlvavhlavhghweepedeekfgaadsglveeprnas 120  
Qy 121 LQAOVVLSPQAYPTARCVLLEQVPAALVQFGSGVYVDCFEAALGSEVRWISYQTPR 180  
Dy 121 lqaovvlsqayptarcvllvqvpaaalvqfgsgsvyvdceaalgsevrilwysyqtp 180  
Qy 181 YKELNHTQOLPALPWLNVASDGNVHLNVSEEHFGLSLVWNOVQGPCKPRWKNLT 240  
Dy 181 ykelnhtqolpalpwlvnsadgdnvhlvlnvseeqhfglslvwnvqgppkprwknlt 240  
Qy 241 GPQITLNLHTDLVPCICIQVWPLEDSVRTNICPFREDPRAHQLNMQAARLLTLQSWL 300  
Dy 241 gpqitlnhtdlvpcicqvwplesdvrtnicpfredprahqnlwqaarlrltlqswl 300  
Qy 301 LDAPCSLPAEALCWAPGDCQPLVPPLSWENVTVDKLEFPLKKGHNLCVQVNSSE 360  
Dy 301 ldapcslpaealcwapgdcqplvpplswenvtvd-----vnsse 360  
Qy 361 KQLQECIWLADSLGPKDDVLLLETRGPQDNRSLCALEPSGCTSLPSKASTRAARLGEYL 420  
Dy 361 kqlqeciwadslgpkddvllletrgpqdnrslcalepsgctslpskastaararlg 420  
Qy 344 KQLQECIWLADSLGPKDDVLLLETRGPQDNRSLCALEPSGCTSLPSKASTRAARLGEYL 403  
Qy 421 LQDLQSGGQCQLQWDDDLGALWACPMNDKYIHKRWALVWLACLLFAAALLLLKKDHAK - 479

Db 404 lqdlqsgqcqlwdddlgalwacpmdkyihkrwalvwlacilfaaalslillkkdhak 463  
Qy 480 -----AAAGRAALLYSADDSGFERLVGALASALCOLPLRVAVDLNSRRELS 527  
Dy 464 wlrllkqdvrsagaaragaaalllysaddsgferlvgalasalcolplrvavdlwsrrels 523  
Qy 528 AOGPVAMFHAQRROTLOEGGVVLLFSPGAVALCSEMLQDGVSGGAHGHDAFRASLSC 587  
Dy 524 aggpvawfhaqrrotlqegggvvllfsgpavalcsewllqdgvsppgaingphdafasrlsc 583  
Qy 588 VLPDFLQGRAPGSYVGACDFDRLLHPDAVPALFRTVPVFTLPSQLPDFLQALQQRAPRSG 647  
Dy 584 vlpdfqlgrapgsyvgacdfdrllhpdavpalfrtvpvftlpsqpdfqlgalqqrprsg 643  
Qy 648 RLQERAEOVSRAQLPALDSYFHPHPTGTPAPGRGVGPGAGPGAGDGT 692  
Dy 644 rlqeraeovsralqpalidsyfhpptpapgrrgvpgagpgagdgdt 688

RESULT 8  
AAB61881  
ID AAB61881 standard; Protein; 575 AA.  
XX  
AC AAB61881;  
XX  
DT 08-MAY-2001 (first entry)  
XX  
DE Human variant zcytor14 protein zcytor14-1.  
XX  
KW Cytokine receptor; zcytor14; human; inflammation; rheumatoid arthritis;  
KW antiinflammatory; gene therapy; vaccine; variant; zcytor14-1.  
XX  
OS Homo sapiens.  
XX  
PN WO200104304-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 30-JUN-2000; 2000WO-US18383.  
XX  
PR 07-JUL-1999; 99US-0348854.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Burkhead SK, Pownder SL;  
XX  
DR WPI; 2001-112618/12.  
XX  
PT N-PSDB; AAC85029.  
XX  
PT New polypeptide encoding a human cytokine receptor zcytor14, for  
XX treating inflammation e.g. rheumatoid arthritis -  
PS Disclosure; Page 2-3; 112pp; English.  
XX  
CC The invention provides a new human cytokine receptor designated zcytor14.  
CC zcytor14 can be expressed by standard recombinant methodology. The  
CC encoding nucleic acid is useful for detecting the expression of a  
CC zcytor14 gene in a biological sample. Anti-zcytor14 antibodies can be  
CC used to screen biological samples in vitro for the presence of zcytor14.  
CC proteins, polypeptides and peptides having zcytor14 activity can be  
CC administered to a subject who lacks an adequate amount of this  
CC polypeptide, for treating inflammation and conditions such as rheumatoid  
CC arthritis. In contrast, zcytor14 antagonists (e.g. anti-zcytor14  
CC antibodies) can be used to treat a subject who produces an excess of  
CC zcytor14. zcytor14 nucleotide sequences can also be used to provide  
CC zcytor14 to a subject. The present sequence represents a variant of  
CC the human cytokine receptor zcytor14, designated zcytor14-1. This  
CC variant is a truncated form of the receptor polypeptide and lacks  
CC amino acid residues 1-113 of zcytor14.  
XX  
SQ Sequence 575 AA;

Query Match	80.0%;	Score 2971;	DB 22;	Length 575;
Best Local Similarity	94.9%;	Pred. No. 9.6e-263;		
Matches 562;	Conservative 0;	Mismatches 0;	Indels 30;	Gaps 2;
QY 114	EEPRNASLQAVVLSFQAPRTARCVLLEVOVPALVFGQSVGVYDCEFAALGSEVRI	173		
Db 1	eeprnaslqavvlsfqaiprtarcvilevqpaalvfgqsvgvdydcefaalgsevril	60		
QY 174	WSYTPRYKEKLNHTQOLPALPWLNSADGDNVHLVNLNVEEOHFGLSLYWNOVGPPKP	233		
Db 61	wsytprykelnhtqolpalpwlwlnsvadgdnvhlvlnvseeqnfglislywnvgppkp	120		
QY 234	RHKNLTGPQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAHONLMOAARLRL	293		
Db 121	rwhknltgpqiitlnhtdlvpcicqvwplepdsrvtrnicpfredprahonlmoaarlr	180		
QY 294	LTLQSWLLDAPCSLPAEALCWAPGDDPCQPLVPPLSWENVTVDKVLFFPLKLGHNLC	353		
Db 181	ltlqswlldapcslpaealcwragddpcqplvpplswenvtvd-----	225		
QY 354	VQVNSSEKLQLOECIWDLSGLPKDDVLLLETRGPDDNRSLCALPESGCTSLPSKASTRA	413		
Db 226	--vnsseklqlceclwadslgplkddvllletrgpddnrslcalepsegctslpskasta	283		
QY 414	ARLGEYLLQDLSQGCQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILL	473		
Db 284	arlgeyllqdlsggcqlqlddldgalwacpmdkylhkrwalvwlacllfaaalsllll	343		
QY 474	KDHAKE-----AAAGRAALLYSADDSGFERLVGALASALCOLPURAVDL	520		
Db 344	kdhakewlllkqdvrsaaagraalllysaddsgferlvgalasalcqlplravadl	403		
QY 521	WSRRELSSAGPVAVFHAQRRQTLQEGGVVLLFSPGAVALCSEWLODGVSGPAGHPHDA	580		
Db 404	wsrrelssagpvavfhaqrtrtlqeggvvllfsgpavalcsewldgvgsgpangphda	463		
QY 581	FRASLSCVLPDFLQGRAGSYVGACFDRLHPDAVPALEFRTVPVFTLPSQLPDLFALQ	640		
Db 464	fraslscvlpdflqgragsyvgacfdrlhpdavpalftrvpvftlpqlpdlfalgq	523		
QY 641	PRAPSGRLQRAEQVSRLQPALDSYFHPHGGTPAPGRGVGPGAGGAGDGT	692		
Db 524	prapsgrlqraeqvsrlqpaldsyfhphtgtpapgrgvpgagagdgdt	575		
RESULT 9				
AAV76143				
ID	AAV76143 standard; Protein; 309 AA.			
XX				
AC	AAV76143;			
XX				
DT	23-MAR-2000 (first entry)			
XX				
DE	Human secreted protein encoded by gene 20.			
XX				
KW	Human; secreted protein; cancer; tumour; developmental abnormality;			
KW	foetal deficiency; blood disorder; immune system disorder; inflammation;			
KW	autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;			
KW	schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;			
KW	atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;			
KW	digestive disorder; endocrine disorder; infection; AIDS; leukaemia;			
XX	therapy; chromosome 3.			
OS	Homo sapiens.			
XX				
PN	W09958660-A1.			
XX				
PD	18-NOV-1999.			
XX				
PF	06-MAY-1999; 99MO-US09847.			
XX				
PR	12-MAY-1998; 98US-0085093.			

PR 12-MAY-1998;	98US-0085094.
PR 12-MAY-1998;	98US-0085105.
PR 12-MAY-1998;	98US-0085180.
PR 18-MAY-1998;	98US-0085906.
PR 18-MAY-1998;	98US-0085920.
PR 18-MAY-1998;	98US-0085921.
PR 18-MAY-1998;	98US-0085922.
PR 18-MAY-1998;	98US-0085923.
PR 18-MAY-1998;	98US-0085924.
PR 18-MAY-1998;	98US-0085928.
PR 18-MAY-1998;	98US-0085925.
PR 18-MAY-1998;	98US-0085927.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
PI	Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
PI	Lafleur DM, Endress GA, Ebner R;
XX	
DR	WPI; 2000-062296/05.
DR	N-PSDB; AA265269.
XX	
PT	New isolated human genes and the secreted polypeptides they encode,
PT	useful for diagnosis and treatment of e.g. cancers, neurological
PT	disorders, immune diseases, inflammation or blood disorders -
PS	
XX	
XX	Claim 11; Page 372-373; 475pp; English.
CC	AA265250 to AA265350 represent 97 isolated human secreted protein genes.
CC	AAV76124 to AAV76223 are the secreted proteins encoded by the 97 human
CC	genes. The gene encoding this protein was found to be on chromosome 3.
CC	The genes and their corresponding secreted polypeptides are
CC	useful for preventing, treating or ameliorating medical conditions,
CC	e.g. by protein or gene therapy. Also pathological conditions can be
CC	diagnosed by determining the amount of the new polypeptides in a sample
CC	or by determining the presence of mutations in the new genes. Specific
CC	uses are described for each of the 97 genes, based on which tissues they
CC	are most highly expressed in, and include developing products for the
CC	diagnosis or treatment of cancer, tumours, developmental abnormalities
CC	and foetal deficiencies, blood disorders, diseases of the immune system,
CC	autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC	disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC	disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC	disorders, digestive/endocrine disorders, infections and AIDS. The
CC	polypeptides are also useful for identifying their binding partners.
CC	The sequences shown in AAV76224 to AAV76424 represent fragments of the
CC	secreted proteins.
XX	
SQ	Sequence 309 AA;

Query Match	38.5%;	Score 1432;	DB 21;	Length 309;
Best Local Similarity	90.9%;	Pred. No. 2.4e-122;		
Matches 270;	Conservative 0;	Mismatches 11;	Indels 16;	Gaps 2;
QY 1	MPVPWFLLSLALGRSPVVLRLVGPDPDATHCSPGLSCLRLWSDILCLPGDIVPAGPV	60		
Db 1	mpvpwfllslalgrspvvlrlvvgpdathcspglscrlwdsdillclpgdivpapgvp	60		
QY 61	LAPTHLQTELVLRCQKEDCDLCLRVAVHLVAVHGHWEPEDEKFGGAADSGVEPRNAS	120		
Db 61	laphlqltelvlrcqketcldclrvxhvlavhghwepeedeekfggaadlgveeprnas	120		
QY 121	LQAOVLVSFQAYPTARCVLLEVPALVQFGQSVGVYDCEFAALGSEVRINWYTQPR	180		
Db 121	lqavvlvsfqaiprtarcvilevqpaalvfgqsvgvdydcefaalgsevrinwytqpr	180		
QY 181	YEKELNHTQQLP-----ALPWLNVSAODGNVHLVNLNVEEOHFGLSLYWN	225		
Db 181	yekelnhtqqipdclrglevwnslpwcwlpwlnvsadgdnvhlvlnvseeqnfglislywn	240		
QY 226	QVQGPAPKPRWHKNTLGTQIITLNTDLVPCICIQWPLEPDSVRTNICPFREDPRAH	282		
XX				

Db 241 qvqppkprwhkntgqitlnhtdlvpcicqvwpdepdvrtssapgrtp-ah 296

RESULT 10  
AAB88448  
ID AAB88448 standard; Protein; 267 AA.  
XX  
AC AAB88448;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Human membrane or secretory protein clone PSEC0233.  
XX  
KW Human; secretory protein; membrane protein; vaccine; gene therapy;  
KW rheumatoid arthritis; diabetes.  
XX  
OS Homo sapiens.  
XX  
PN EPI067182-A2.  
XX  
PD 10-JAN-2001.  
XX  
PF 07-JUL-2000; 2000EP-0114090.  
XX  
PR 08-JUL-1999; 99JP-0194179.  
PR 11-JAN-2000; 2000JP-0118775.  
PR 02-MAY-2000; 2000JP-0183766.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;  
XX  
DR WPI; 2001-093989/11.  
DR N-PSDB; AAF93875.  
XX  
PT Nucleic acids encoding secretory proteins/membrane proteins, useful in  
PT gene therapy or as candidate target molecules in drug development -  
XX  
PS Claim 1; SEQ ID 264; 609pp + CD ROM; English.  
XX  
CC This invention relates to nucleic acid sequences AAF93744 - AAF93916  
CC which encode human secretory or membrane proteins represented by  
CC AAB88317 - AAB88419. Included in the invention are primers  
CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the  
CC cDNA sequences of the invention. The invention also includes methods for  
CC the production of antibodies directed against the proteins, and cDNA  
CC sequences, which can be used in vaccines. The polynucleotide sequences  
CC can be used in gene therapy. The polynucleotide sequences and the  
CC proteins they encode may be used in the prevention, treatment and  
CC diagnosis of diseases associated with inappropriate secretory  
CC protein/membrane protein expression. The nucleic acids and complementary  
CC sequences may also be used as DNA probes in diagnostic assays  
CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the  
CC presence of similar nucleic acid sequences in samples. They may also be  
CC used to study the expression and function of secretory proteins/membrane  
CC polypeptides and their role in metabolism. The polypeptides may be used  
CC as antigens in the production of antibodies against them and in assays to  
CC identify modulators (agonists and antagonists) of expression and  
CC activity. The antibodies and antagonists may also be used as therapeutic  
CC agents to down regulate expression and activity. The antibodies may also  
CC be used as diagnostic agents for detecting the presence of the  
CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay  
CC (ELISA)). Examples of diseases which may be treated include rheumatoid  
CC arthritis and diabetes.  
XX  
SQ Sequence 267 AA;

Query Match 35.1%; Score 1305; DB 22; Length 267;  
Best Local Similarity 89.2%; Pred. No. 8e-111;  
Matches 249; Conservative 1; Mismatches 17; Indels 12; Gaps 2;

Qy 1 MPVPWFLLSLALGRSPVVLRLVGPQDATHCSPGLSCLRLWDSIDLGLPGDIVPAPGPV 60

Db 1 mpvpwfllslalgrspvvlrlvqpqathcsplscrlwdsdlclpgdivpapg 60  
QY 61 LAPTHIQTELVLRCQKETDCDCLRLVAVHVLAVHGHWEPEDEKFGAAGSVEEPRNAS 120  
Db 61 lapthiqtelvlrcqketdcdclrlvavhvlavhghwepeedeekfgaadsgveeprnas 120  
QY 121 LQAQVVLSTQAYPTARCVCVLELVQVPAALVQFGSGVSVVYDCFEAALGSEVRWISYTPR 180  
Db 121 lqaqvvlstqayptarcvcvllievqvpaaalvqfgsgsvsvvdydcfeaalgservriwstqpr 180  
QY 181 YEKELNHTQOLPALPWLNVSDGDNVHLVNLVSEEHFGLSLYVNOYQGGPKPRWHKNTL 240  
Db 181 yekelnhtqqlpalpwlinvdsadgdnvhlvlnvseeqhfglslvynvqvggpkprwhknlv 240  
QY 241 GPOIITLNLHTDLVPCLCIQVMPLEPDSVRTNICPFREDP 279  
Db 241 rpp-----psqvshsc-----rmpvpvqrtqchiredp 267

RESULT 11  
AAM25864  
ID AAM25864 standard; Protein; 332 AA.  
XX  
AC AAM25864;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:1379.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antitaggagant; haemostatic; vulnerary; antitumor; osteopathic; eczema;  
KW dermatological; anti-allergic; antiasthmatic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US35017.  
XX  
PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-457603/49.  
XX  
DR N-PSDB; AAM99805.  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
PS Claim 20; Page 283; 1217pp; English.  
XX  
CC AAM99166 to AAM99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;

PN	WO200166722-A1.
XX	
XX	
PD	13-SEP-2001.
XX	
PF	07-MAR-2001; 2001WO-US07260.
XX	
XX	
PR	08-MAR-2000; 2000US-187873P.
PR	11-AUG-2000; 2000US-224367P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI	Gruber JR, Endress GA, Ruben SM;
XX	
XX	WPI: 2001-589939/66.
DR	N-PSDB: AAD18276.
XX	
PT	Novel isolated immune system-related polypeptide useful for treating
PT	rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT	diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT	viral hepatitis -

Claim 11; Page 312-313; 315pp; English.

The invention relates to human immune system-related protein and their DNA. Human immune-system related protein and DNA are useful for preventing, treating or ameliorating a medical condition in a mammalian subject, for diagnosing, preventing or treating immune system-associated disorders, autoimmune disorders (rheumatoid arthritis), inflammatory disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders (allergies), infectious diseases (e.g., viral hepatitis), complement activation disorders, immune complex diseases, neoplastic disorders (cancer), hyperproliferative disorders (Caucher's disease), disorders associated with neovascularisation, diseases at the cellular level, cardiovascular disorders (arrhythmias), wound healing and epithelial cell proliferation, endocrine disorders (diabetes mellitus) and neurological disorders (ischaemic lesions). Immune-system related protein or DNA is useful for preventing hair loss, skin aging due to sunburn, to maintain organs before transplantation, to treat weight disorders, to modulate mammalian characteristics, to change a mammal's mental or physical state, or as a food additive or preservative. Immune-system related DNA is useful in gene therapy, for chromosome identification, radiation hybrid mapping, long range restriction mapping and in forensic biology. The present sequence represents a human immune-system related protein of the invention.

```

SQ Sequence 204 AA;
      Query Match          29.0%; Score 1079; DB 22; Length 204;
      Best Local Similarity 100.0%; Pred. No. 2.5e-90;
      Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 489 LLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFWFIHAQRQTLOEGGV 548
      |||||
      1 llysaddsgferlvgalasalcqiplrvavdlwsrreelsagpgpvafwfihaqrrgtlqeggv 60
      |||||

QY 549 VVLLFSPGAVALCSEWLQDGVSGPGAGHPHDAFRASLSCVLPDFLQGRAPGSGVYGACFDR 608
      |||||
      61 vvllfspgavalcsewllqdvsgpgaghhphdafraslscvlpdfiqgrapgsyvgacfr 120
      |||||

QY 609 LLHPDVPALFRTVPVFTLLPSQLPDFLGLALCQOPRPSGRQLQERAEQVSRALQPALDSYF 668
      |||||
      121 llhpdvpalfrtvpvftllpsqlpdfilgalqgrprsgriqeraeqvsraiqpaldsyf 180
      |||||

QY 669 HPDGTAPAGRGVPGAGPGAGDGT 692
      |||||

Db 181 hppgtapagrvgpgagpgagdgct 204
      |||||

RESULT 13
AAV76048

```

XX AC AAY76048;  
 XX DT 27-MAR-2000 (first entry)  
 XX DE Murine skin cell protein, SEQ ID NO:303.  
 XX KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 KW secreted; transmembrane; inflammation; cancer; neurological disease;  
 KW angiogenesis; tumour vascularisation; growth disorder;  
 KW developmental disorder; skin wound; hair follicle disorder;  
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.  
 XX OS Mus sp.  
 XX FN WO9955865-A1.  
 XX PD 04-NOV-1999.  
 XX PF 29-APR-1999; 99WO-NZ00051.  
 XX PR 29-APR-1998; 98US-0069726.  
 XX PR 09-NOV-1998; 98US-0188930.  
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;  
 XX WPI; 2000-072177/06.  
 XX DR N-PSDB; AAZ61753.  
 XX PT Novel polynucleotides useful for the treatment of various conditions  
 XX including wounds and cancer -  
 XX PS Claim 4; Page 179-180; 235pp; English.  
 XX CC The invention relates to novel nucleic acid sequences derived from rat  
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
 CC cells. Polypeptides of the invention may be used to treat inflammation,  
 CC cancer and neurological diseases. The proteins may be used to stimulate  
 CC the growth and motility of keratinocytes, to inhibit the growth of  
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to  
 CC modulate skin inflammation, to modulate epithelial cell growth and to  
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used  
 CC to treat growth and developmental defects, skin wounds and hair follicle  
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded  
 CC by cDNA sequences derived from several mouse, rat or human skin cell  
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and  
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating  
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,  
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more  
 CC putative transmembrane domains.  
 XX SQ Sequence 617 AA;  
 Query Match 8.8%; Score 326.5; DB 21; Length 617;  
 Best Local Similarity 26.7%; Pred. No. 8.7e-21;  
 Matches 132; Conservative 56; Mismatches 184; Indels 123; Gaps 19;  
 QY 237 KNLGTGQIITLHNTDLVPCICIQWVPLEPDSVRTNICPFREDPRAH-QNLWQAR----- 290  
 DB 183 kivsghvtldpyefilpmcieceasyldqetvrrkkcpqgsweagysdfwqsirfdys 242  
 QY 291 -----LRLTLQSWLLDAPCSLPAAALCWAPGPGPCQPLVPLSWENTVTVKVFEPFL 345  
 DB 243 qhngvmaltlr-----cpkleaslcwrgdplpcetl-----pnataqesegwyi 289  
 QY 346 LAG---HPNLCVQVNSSEKIQEQELWADSLGPKDDVLLLETRGPQDNRSKALEPSGC 402  
 DB 290 lenvdhlpqlcfkf-sfensshvec--phqsgslpswtvsmtdtgaq-----ltlhfsr 341

QY 403 TSLPSKASTRAARLG-----EYLLQDLQSGCLOLWDDLGAL 440  
 DB 342 tyatfsaawsdpglpgdtpmpvpyisqtgsvptlilipfrqencilivrsdvhfa 401  
 QY 441 W---ACPMKYIHKRWALVWLACLLFAAALSILILKKDHAKAAAGRAALLLYSADDSG 497  
 DB 402 wkhvlpdd-----apyptqllr-----slsggrtpvlllhaadsea 440  
 QY 498 FERLVGALASALCOLPRLVA-----VDLWSRRELQAQGPVAFWFAORRQTLQEGGV 549  
 DB 441 qrlivgalael-----lrtalggdrvldlwegthvarigplwlaarervaregtv 495  
 QY 550 VLFSPGAVALCSEWLQDGVSGPAGHGFDAFRASLSCVLPDFLOGRAPGVSYGACFRL 609  
 DB 496 lllwn-----cagpstacsgdpqaasrltil-----caaprplllayfsl 536  
 QY 610 LHPDAVPALFRTVPVETLPSOLPDFLGALQ-OPRAPRSG-----RLQRAEOVS-RA 659  
 DB 537 cakgdiprprlprlrrldprllraldaqpatlasswshlgakrcldknrlqechlle 596  
 QY 660 LQPALDSYFHPPGTP 674  
 DB 597 leaakddyqgstnsp 611  
 RESULT 14  
 AAB55987  
 ID AAB55987 standard; Protein; 617 AA.  
 AC AAB55987;  
 XX 08-MAR-2001 (first entry)  
 XX Skin cell protein, SEQ ID NO: 303.  
 KW Mouse; skin cell; cytostatic; antiinflammatory; anti-HIV;  
 KW neurotropic; neuroprotective; vulnery; immunomodulatory; vaccine;  
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;  
 KW inflammation; neurological disease.  
 OS Mus sp.  
 XX WO200069884-A2.  
 XX 23-NOV-2000.  
 XX 15-MAY-2000; 2000WO-NZ00075.  
 XX 14-MAY-1999; 99US-0312283.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;  
 XX WPI; 2001-007495/01.  
 XX N-PSDB; AAC99686.  
 XX New isolated polynucleotide used in the identification of genetic  
 XX disorders and encoding polypeptides used for treating inflammatory  
 XX disease, cancer and neurological diseases -  
 XX Claim 4; Page 244-245; 352pp; English.  
 XX The present sequence is a polypeptide which is expressed in  
 XX mammalian skin cells. The polypeptide is useful for stimulating  
 XX keratinocyte growth and motility, inhibiting the growth of cancer cells,  
 XX modulating angiogenesis, inhibiting angiogenesis and vascularisation of  
 XX tumours, modulating skin inflammation, stimulating the growth of  
 XX epithelial cells, inhibiting the binding of human immunodeficiency virus  
 XX (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and  
 XX neurological diseases. The polynucleotide can be used as a marker, in  
 XX the identification of genetic disorders, and for the design of



CC oligonucleotides for examining expression patterns.

XX Sequence 617 AA;

Query Match 8.8%; Score 326.5; DB 22; Length 617;

Best Local Similarity 26.7%; Pred. No. 8.7e-21;

Matches 132; Conservative 56; Mismatches 184; Indels 123; Gaps 19;

QY 237 KNTGQITITLNTDLPCLICIQVPLEPDSVRTNICPFREDPRAH-QNLWQAAR----- 290  
 Db 183 kivsghgtvdlpyellfpcmcieasyldetvrrkkcpfsgpeaygsdfwqsfirftds 242  
 QY 291 -----LRLTLQSWLLDAPCSLPAEALCWRAFGDPCQPLVPLSWENVTVDKVLPEPL 345  
 Db 243 qhnmqvmaltlr-----cplkleaslwrqdpptcetl-----pnataqesegwyi 289  
 QY 346 LKG---HPNLVCQVNSEKIQLOECLWADSLGPKDDVLLLETRGPQDNRSLCALEPSGC 402  
 Db 290 lenvdhlpqlcfkf-sfensshvec--phgsgslpwtvsmtdtaqg-----ltlhfsr 341  
 QY 403 TSLPSKASTRAARLG-----EYLLOLQSGCQLQLWDDDLGAL 440  
 Db 342 tyatfsaawsdgplgdpdpmpvysisqtgsvptvtdlilpflrqencilwrsdvhfa 401  
 QY 441 W---ACPMKYIHKRWALVWLACLLFAAALSLLKLDHAKAAARGAALLLYSADDSG 497  
 Db 402 wkhvlecpdd-----apytqlllr---slgsgtrtrpvlllhaadsea 440  
 QY 498 FERLVCALASALCOLPLRVA-----VDLWSRRELSAOGPVANFHAQRQTLOEGGVV 549  
 Db 441 qrrlvgalael-----lrltaggrdvldwethvargipwlaarervaregtv 495  
 QY 550 VLLFSPGAVALCSEWLQDVGSGAHPDFAFRASLSCVLPDFLOGRAPGSYVCACFDRL 609  
 Db 496 illwn-----cagpstacsqdpqaasrltl-----caaprpillayfsrl 536  
 QY 610 LHPDAPALFRTVPVFTLSQLPDFLGLAQ-QPRAPRSG-----RLQERARQVS-RA 659  
 Db 537 cakgdiprpralpryllrdiprllraidaqpattlasswshlgakclknrlqeqchlle 596  
 QY 660 LOPALDSYFHPGTP 674  
 Db 597 leaakddygstnsp 611

# RESULT 15

AAU04957  
 ID AAU04957 standard; Protein; 667 AA.

XX AC AAU04957;

XX DT 24-OCT-2001 (first entry)

XX DE Human Interleukin 17 receptor, IL-17RH3.

XX KW Human; Interleukin-17 receptor; IL-17RH3; agonist; antagonist;

XX KW PRO9877; DNA 119502-2789; systemic lupus erythematosus;

XX KW rheumatoid arthritis; osteoarthritis; diabetes mellitus;

XX KW allergic disease; asthma; demyelinating disease;

XX KW degenerative cartilaginous disorder; transplantation associated disease.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..23

FT FT /label= Signal\_peptide

FT FT 24..667

FT FT /label= Mature\_IL\_17RH3

FT FT 90..96

FT FT /note= "N-myristoylation site"

FT FT 104..108

FT FT /note= "cAMP/GMP-dependent protein kinase

FT Modified-site phosphorylation site"  
 FT 318..322 /note= "Asn is N-glycosylated"  
 FT 322..329 /note= "Tyrosine kinase phosphorylation site"  
 FT 347..351 /note= "Asn is N-glycosylated"  
 FT 358..364 /note= "N-myristoylation site"  
 FT 364..368 /note= "Asn is N-glycosylated"  
 FT 453..462 /note= "Eukaryotic cobalamin-binding protein motif"  
 FT 455..472 /note= "Transmembrane domain"  
 FT 470..476 /note= "N-myristoylation site"  
 FT 482..486 /note= "Glycosaminoglycan attachment site"  
 FT 645..649 /note= "cAMP/GMP-dependent protein kinase phosphorylation site"  
 FT 645..649 /note= "cAMP/GMP-dependent protein kinase phosphorylation site"  
 FT 645..649 /note= "cAMP/GMP-dependent protein kinase phosphorylation site"  
 XX XX  
 PN WO200146420-A2.  
 XX 28-JUN-2001.  
 XX 20-DEC-2000; 2000WO-US34956.  
 XX 23-DEC-1999; 99US-0172096.  
 PR 30-DEC-1999; 99WO-US31274.  
 PR 11-JAN-2000; 2000US-0175481.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 02-MAR-2000; 2000WO-US05841.  
 PR 21-MAR-2000; 2000US-0191007.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 22-JUN-2000; 2000US-0213087.  
 PR 22-AUG-2000; 2000US-0644848.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 24-OCT-2000; 2000US-0242837.  
 PR 10-NOV-2000; 2000WO-US30873.  
 PR 28-NOV-2000; 2000US-0253646.  
 PR 01-DEC-2000; 2000WO-US32678.  
 XX XX  
 PA (GETH ) GENENTECH INC.  
 XX XX  
 PI Chen J, Filvaroff E, Fong S, Goddard A, Godowski PJ, Grimaldi CJ;  
 PI Gurney AL, Li H, Hillan KJ, Tumas D, Van Lookeren M, Vandlen RL;  
 PI Watanabe CK, Williams PM, Wood WI, Yansura DG;  
 XX XX  
 DR WPI; 2001-451708/48.  
 DR N-PSDB; AAS09516.  
 XX XX  
 PT Novel PRO polypeptides homologous to interleukin-17, useful for the  
 PT diagnosis and treatment of immune related disease e.g. rheumatoid  
 PT arthritis and diabetes  
 XX XX  
 PS Claim 10; Fig 16; 188pp; English.  
 XX XX  
 CC The sequence is PRO9877 which is the human Interleukin 17 receptor,  
 CC IL-17RH3, encoded by DNA 119502-2789. A composition  
 CC containing ant/agonists to the PRO polypeptides or individual components  
 CC are useful for treating a mammal with an immune related disease, e.g.  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
 CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal  
 CC disease, a demyelinating disease, an autoimmune or immune-mediated skin  
 CC disease, contact dermatitis, an allergic disease e.g. food  
 CC hypersensitivity, asthma, a transplantation associated disease, or a  
 CC chronic inflammatory demyelinating polyneuropathy. Treating a

CC degenerative cartilaginous disorder comprises administering a PRO1031 or  
 CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous  
 CC examples of the diseases and disorders are given in the specification.

XX

SQ Sequence 667 AA;

```

Query Match      8.3%; Score 309.5; DB 22; Length 667;
Best Local Similarity 29.0%; Pred. No. 3.5e-19;
Matches 138; Conservative 54; Mismatches 181; Indels 103; Gaps 23;

Qy 237 KNLGPGQIITLNNHTDLPVPCICIQVWPDPEDSVRTNICPFREDPRAH-QNLWQAAAR----- 290
Db 223 kivsghgtvelpfelipclcieasylqedtvrkkcpqswpeaygsdfwksvhtdys 282
Qy 291 -----LRLTLQSWLLDAPCSLPAEALCWAPGPGCPQPLVPPLSWENVTVDK-----V 340
Db 283 qhtqmvmltr-----cpkleaalqrhdwhlckdl-----pnataresdgwv 329
Qy 341 LEFPLKGNHPLCVQV---NSSEKIQLOECLWADSLGPKDDVLLLETRGPQ-----DNR 392
Db 330 lekvdI--hpqlcfkfsfgnssh---vec--phqtgsitwnvsmtdgaqllilhfssr 381
Qy 393 SLCALEPSCGCTSLP-----SKASTRAARLGEYLLQDLQSGOCLOLWDDDLG 438
Db 382 mhatf--saawslpqlgqdtlvppvytvsgargsspsvldliipfipgpcvivrsvdq 439
Qy 439 ALW---ACPMKDYIHKRWALVWLACILFAAALSILLLK-KDHAKAAARGRAALLLYSAD 494
Db 440 fawkhlilcpdvsvyrh--lgllilallliltlgvialtcrppsgpgparpvlllhaad 497
Qy 495 DSGFERLYGALASALCQLPLRVA-----VDLWSRRELSAQGPVAFHQAQRRTIQEG 546
Db 498 seaqrllvgalael-----lraalggrdvivdlwegrhvarvgpplwlaartvareq 552
Qy 547 GVVVLLFSPGVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAGCS-VYGAC 605
Db 553 gtvlllwsqadlr-----pvsqp-----dpraaplall-----haaprlllilay 593
Qy 606 FDRLLHPDAVPALFRTVPVFTLPSQLPDLFGALQ-QPRAPRS--GRLQERAEQVSR 658
Db 594 fsrlcakgdippprlprllrdlprllraldarpfaeatswgrlqarqrqsr 649
  
```

Search completed: May 19, 2002, 13:26:40  
 Job time: 7621 sec